Vol. 40, No. 4 April, 19

Psychological Bulletin

Social Science
BF
1
.P9
V.40
no.4

EDITED BY
JOHN E. ANDERSON, UNIVERSITY OF MINNESOTA

WITH THE CO-OPERATION OF
S. H. BRITT, GEORGE WASHINGTON UNIVERSITY; W. T. HERON, UNIVERSITY OF MINNESOTA; W. A. HUNT, WHEATON COLLEGE; J. G. JENKINS, UNIVERSITY OF MARYLAND; A. W. MELTON, UNIVERSITY OF MISSOURI; J. T. METCALF, UNIVERSITY OF VERMONT.

CONTENTS

General Review and Summary:

The Analysis of Variance in Psychological Research: HENRY E. GARRETT and JOSEPH ZUBIN, 233

Psychology and the War:

Personnel Research in the Army. III. Some Factors Affecting Research in the Army: STAFF, PERSONNEL RESEARCH SECTION, CLASSIFICATION AND REPLACEMENT BRANCH, AGO, 271

The Special Training Units of the Army: MORTON A. SEIDENFELD, 279

The Military Use of the Rail-Walking Test as an Index of Locomotor Coordination: S. ROY HEATH, JR., 282

A Course in Army Personnel Techniques: CALVIN S. HALL, 285

The Committee on Food Habits: MARGARET MEAD, 290

Attitudes of Youths of High School Age Toward the War: MANDEL SHERMAN, 294

Psychology and the War: Notes, 300

Book Reviews: 305

Books and Materials Received: 309

Notes and News: 310

PUBLISHED MONTHLY (EXCEPT AUGUST AND SEPTEMBER)
BY THE AMERICAN PSYCHOLOGICAL ASSOCIATION, INC.
NORTHWESTERN UNIVERSITY, EVANSTON, ILLINOIS

Entered as second-class mail matter at the post office at Evanston, Illinois, under the Act of March 3, 1879. Additional entry at the post office at Menasha, Wisconsin.

UNIVERSITY OF MICHIGAN LIBRARIES

PUBLIC

THE AMERICAN PSYCH ASSO N

WILLARD L. VALENT

PSYCHOLOGIC

HERBERT S. LA

Princeton U

contains original contributions only, appears

September, and November, the six numbers co

Subscription: $7.50 (Foreign,

PSYCHOLOGIC

JOHN E. ANDR

University of

Contains critical reviews of books and articles

notices and announcements. Appears monthly

about 501 issues. Special issues of the BULLETI

in some department of psychology.

Subscription: $7.00 (Foreign,

JOURNAL OF EXPERI CHOLOGY

G. W. PEART

University of P

Contains original contributions of an experim Appears monthly, four

January, 1936, two volumes per year, each vol numbers constituting about 625

issues. Subscription: $14.00 ($7.00 per volume; F Sample copy: $2.25

PSYCHOLOGICAL

WALTER S. HUN

Brown Univ

Appears monthly, the twelve numbers and an in making volumes of about

500 pages. The journal is devoted to the publica the world's

literature in psychology and closely related fi

Subscription: $7.00 (Foreign,

PSYCHOLOGICAL

JOHN F. DASH

University of Nor

Consists of longer researches or treatises of which it is

impracticable to publish promptly and in units according

to these aims. The publications appear at numbers until

volumes of about 500 pages.

Subscription: $6.00 per v

JOURNAL OF ABNORMAL PSYCH

GORDON W. ALL

Harvard Uni

Appears quarterly, January, April, July, October numbers constituting a volume

of 500 pages. The journal contains original contri field of abnormal and social

psychology, reviews, and notes.

Subscription: $5.00 (Foreign, $5.2

JOURNAL OF APPLIE OGY

DONALD G. PATE

University of

Covers the application of psychology to appears bi-

monthly: January, March, May, July, September, Nov

Subscription: $5.00

Subscriptions, orders, and business

THE AMERICAN PSYCHOLO ASSOCIATION, INC.

NORTHWESTERN UNIVERSI

Social Science
Library
BF
1
.P9
v.40
no.4

Vol. 40, No. 4 April, 1943

Psychological Bulletin

THE ANALYSIS OF VARIANCE IN PSYCHOLOGICAL RESEARCH*

BY HENRY E. GARRETT AND JOSEPH ZUBIN
Columbia University
and
New York State Psychiatric Institute and Hospital

The statistical techniques subsumed under the general term analysis of variance, and associated with the name of R. A. Fisher (**13, 15**), have for some time been employed in agriculture, biology, and related fields. Fisher's book, "Statistical Methods for Research Workers," first published in 1925, has gone through seven editions; while its companion volume, "The Design of Experiments," first published in 1935, is now in its second edition. Fisher's methods were, naturally enough, first taken up in this country by those primarily interested in agricultural research. Several excellent books have appeared in which Fisher's designs are described and illustrated in this general field (**18, 30, 39**).

Analysis of variance has not, as yet, been widely used in psychological research (**3**) or in the social sciences (**9**). Even as recently as 1940, the author of a textbook (**26**) in which Fisher's methods are applied to problems in educational research, found it necessary to use artificial data in several instances because of the lack of experimental material in this field. The neglect of Fisher by psychologists has been owing in part to the difficulty which many encounter in following Fisher's treatment. Moreover, the belief that Fisher's methods are concerned chiefly with small samples has doubtlessly had its influence; as has also the difficulty which arises in readily translating experimental designs dealing with soil fertility, weights of pigs, effectiveness of manurial treatments and the like over into a psychological framework.

* Garrett is responsible for Sections I, II, III, V and VI; Zubin for Sections IV and VII. Dr. John W. Fertig read the manuscript and gave valuable suggestions.

The present review attempts to outline and explain the experimental designs in which analyses of variance* are useful. A minimum of theory has been introduced, as theory and methods of calculation are readily available in several textbooks. What we have tried to do is to cite experiments in which Fisher's experimental designs are used; to explain why they are valuable; and to indicate some of the situations in which we believe analysis of variance techniques offer a more efficient method of attacking the problem than those currently in use.

I. Meaning of Analysis of Variance

In experimental work in which many measures (*e.g.*, scores) have been obtained from individuals or groups, it is often important to be able to separate out or differentiate the factors which contribute to individual or group differences. The measure of variation analyzed is usually the variance, or σ^2, instead of the σ. A simple illustration will show why σ^2 is used. Let Z be a variable which is composed of two components X and Y. X and Y are indicative, say, of masculinity-femininity characteristics; men tend to have higher scores in X, women higher scores in Y. For any individual, Z is the sum of his X and Y scores. Then $Z=X+Y$, or $z=x+y$ when measures are taken from their respective means. Squaring, summing, and dividing by $N-1$ we have

$$\frac{\sum z^2}{N-1}=\frac{\sum x^2}{N-1}+\frac{\sum y^2}{N-1}+\frac{2\sum xy}{N-1}.$$

Now if X and Y are independent so that $r=.00$,

$$\frac{\sum z^2}{N-1}=\frac{\sum x^2}{N-1}+\frac{\sum y^2}{N-1}$$

and the variance of Z has been analyzed into two parts. The proportion which each variance contributes to differences in Z may now be readily determined. Should X and Y *not* be independent, so that

$$\frac{\sum z^2}{N-1}=\frac{\sum x^2}{N-1}+\frac{\sum y^2}{N-1}+\frac{2\sum xy}{N-1},$$

we have, in addition to the independent contributions of X and Y, their joint contribution through $\sum xy/(N-1)$. This last term is

* Analysis of variance, as a general term, covers also analysis of covariance.

called the *covariance* of X and Y, and its relation to correlation is readily apparent. When comparable groups are combined, the variance of the resulting inclusive group may be broken down into the sum of the variances and covariances of the component groups. Because of the additive nature of variances, they are analyzed instead of the σ's, when one is seeking the contributing causes to group or individual differences.

An instance in which the variance of a final measure is analyzed in terms of its components is Shuttleworth's study (**36**) of the contributions to differences in IQ of heredity, environment, and accidents plus intra-family environmental differences. Let x=IQ, h=heredity, e=environment, and a=accidents, etc.; then if $x=h+e+a$,

$$\sigma_x^2 = \sigma_h^2 + \sigma_e^2 + \sigma_a^2 + 2r_{he}\sigma_h\sigma_e.$$

Going on from here, Shuttleworth analyzes the variations in IQ into the proportions contributed by the factors h, e, a, and he in combination. In the studies of Roff and Roff (**32**) and of Rubin-Rabson (**33**), the total variance of a final result is analyzed into several components. The relative weights of these components are determined from the percentage which each contributes to the total V.* All of the examples cited use fairly simple experimental designs. But no matter how complex the final analysis, the breakdown of the total variance into components and contributing variances is the essential objective of the method.

II. Simple Classification of Variates

The simplest problem in which the analysis of variance approach is often profitable is that in which there are several independent groups of subjects upon all of whom the same measurements have been made. To illustrate, suppose that a test of memory-span for digits has been administered to 5 groups of children (100 in each group) of different national or natio-racial origin, for example, American, Chinese, Mexican, Italian, and Japanese. Our problem is to discover whether these groups differ significantly in memory span. Fisher has shown that the best working hypothesis in situations like this is the "null hypothesis" (**13, 18**); that is, the assumption that there are no "real" differences among our groups. Put somewhat more precisely, the null hypothesis assumes that our 5 samples have been drawn in the same way from

* Variance.

some defined parent population; and that these samples do not differ from one another except through fluctuations which inevitably arise from sampling. The purpose of our experiment may be said to give the facts a chance to disprove this null hypothesis. A null hypothesis, since it is the statement of a negation, can be disproved or refuted but can never be proved. However, if the null hypothesis is *repeatedly* refuted (as for instance in the present illustration), we may finally be led to conclude that there *are* real differences among our groups. But such a generalization—which is the goal of the inductive method—represents a formulation based upon logical consideration as much as upon statistical tests.

To return to our 5 groups, the first step in the analysis of our data is to calculate an over-all or total variance, V_t, from a combination of the scores from all 5 groups. V_t may then be separated into two parts, one portion being the V *within* groups (V_w), the other the V *between* groups (V_b), or among the weighted means of the 5 groups. The calculations (**37**) which involve only simple arithmetic may be tabulated in the following way (data are artificial):

TABLE I

Source of Variation	*Degrees of Freedom*	*Sum of Sqs.*	*Mean Sq. (Variance)*
Between groups	4	8	$2(V_b)$
Within groups	495	990	$2(V_w)$
Total	499	998	$2(V_t)$

Test of Significance, $F = V_b/V_w = 2/2 = 1$

Under "Source of Variation" we place the total V and the two parts which make it up. To obtain the mean square or variance attributable to the "total," "within," and "between" separately, we divide the sum of squares segregated under each category by the DF (degrees of freedom) assigned to it. Finally, to discover whether the mean differences among our five groups are significant, we divide V_b by V_w to get the quantity "F."* The distribution of these F-ratios has been found, and tables (**18, 37**) giving the values of F for different degrees of freedom in numerator and denominator have been calculated at the two critical points, .05 and .01. For an F, the numerator of which has 4 DF and the

* It is possible to test the significance of the "between means" variance in terms of the "within" variance by using Fisher's z-distribution (**18, 30**) instead of the F-distribution. Since $z = \frac{1}{2} \log_e F$, the two tests give the same result. The F-test is somewhat easier to use.

denominator 495 *DF*, the .05 point is 2.39 and .01 point is 3.36. These values may be interpreted to mean that on the null hypothesis (*i.e.*, assumption of no real differences among groups), repetitions of our experiment would 5 times in 100 trials give *F*'s as large as (or larger than) 2.39; and once in 100 trials would give *F*'s as large as (or larger than) 3.36. It is clear that our *F* of 1.00 is not significant, as it is less than the .05 value, and would occur much more often than 5 times in 100. Hence the null hypothesis is *not* disproved, and there is no evidence that the differences among our group means are the result of anything more than sampling fluctuations.

Several points raised by this problem must be considered further. The concept of "degrees of freedom" will be new to many psychologists (**43**). The roots of the notion lie in fairly difficult mathematics, but the rationale of its usage is readily apparent. *DF* refers to the independence or lack of restrictions in the data. If we have 10 scores and their sum, this series possesses only 9 degrees of freedom, since in the case of 9 scores and their sum the tenth is at once determined. If we calculate the mean (which is based, of course, upon the sum) from a distribution, we use up 1 *DF*. In the example above, the *DF* for the total is 500−1 or 499; for between groups, it is 5−1 or 4; for within groups it is 499−4 or 495. The last *DF* may also be determined as follows: since there are 99 *DF* in each group (100−1) and 5 groups, $99 \times 5 = 495$.

The conventional method of attacking a problem in which 5 groups are to be compared for mean performance is to calculate the means and standard errors of the means for each group. The 10 possible comparisons of means are then made [each in terms of its s.e. (difference)] and the critical ratios are evaluated in terms of the probability that such a difference (or a greater) can be attributed to sampling alone. Analysis of variance is superior to the conventional method in several respects. *First*, it should be noted that the standard error (here the within groups *V*) is based upon *all* of the data instead of upon 1/5, and hence is a more stable estimate of experimental error than are the s.e.'s calculated from the separate groups. Throwing our groups together is in accordance with the null hypothesis which we have assumed, namely, that the *M*'s of our 5 groups are essentially equal. To be sure, this null hypothesis may be disproved. But even if it is, the *F*-test is not thereby invalidated, since the net effect of such a finding is to increase *F*—if anything—and hence render the test of significance more sensi-

tive. *Secondly*, the *F*-test provides an over-all test of significance. If it is *not* significant, ordinarily no further test need be made; for none of the mean differences is likely to be significant. If it *is* significant, then the individual *M*'s may be compared and their differences evaluated in the usual way. A significant *F* does not guarantee that *all* mean differences are significant. *Thirdly*, an experimental error based upon the *within* groups variance provides an estimate of the population variance from which differences in group means have been eliminated. Hence, we get a valid estimate of experimental error even when true differences exist among means. *Finally*, note that *all* of the variances in our table, "total," "within" and "between" are the same: namely, 2.00. When the total *V* equals the "within" and the "between," the several groups may be treated as samples drawn from the same population:—the null hypothesis is not refuted. Under these conditions, the three estimates of variance are equally good gauges of experimental error. The discrepancy between the V_b got from the *M*'s and the V_w from within groups is thus a direct measure of the extent to which the means differ.

The simple experimental design described above has been used in several studies in psychology.

Schrader (**34**) has investigated the relation of scores made by college freshmen on a liberalism questionnaire to such variables as best-liked high school subject, mother's education, father's education, father's occupation, and the like. The questionnaire contained 186 statements and covered various types of social problems dealing with labor, race relations, economic relations, and nationalism. There were 149 subjects. Table II shows the data when the men subjects were grouped into 10 categories in accordance with their expressed preferences for high school subjects:

TABLE II

Source of Variation	*DF*	*Sum Sqs.*	*Mean Sq.* (*V*)
Between the 10 categories	9	5.33	.59
Within groups	139	157.21	1.13
Total	148	162.54	

$$F = .59/1.13 = .5 \quad F\begin{cases} .05 = 2.00 \\ .01 = 2.63 \end{cases} \qquad (37)$$

The nonsignificant *F* (.5) gives no evidence of other than a chance relationship between liberalism and best-liked high school subjects. When, however, the masculine subjects were grouped into *two* categories—those whose mothers went to college and those whose mothers did not—the *F*-test for these groups was 9.2, which should be compared with an .01[$F_{.01}$] of 6.8 (obtained from the table of *F*). We must expect, therefore,

an F of 6.8 (or larger) to result from sampling fluctuations once in 100 trials; and accordingly, the F of 9.2 is quite significant of a real difference between the liberalism scores of students whose mothers went to college and those whose mothers did not attend college.

One of the problems attacked by Rietz (**31**) was that of discovering whether students admitted to college from High School *A* over a five year period varied significantly in intelligence from year to year. Intelligence was measured by the American Council Intelligence Examination and was expressed in percentile scores. For 121 students admitted to college over a five year period, we have the following data:

TABLE III

Source of Variation	*DF*	*Sum Sqs.*	*Mean Sq.* (V)
Between the 5 groups	4	236.97	59.24
Within groups	116	570.14	4.92
Total	120	807.11	

$F = 59.24/4.92 = 12.0$ $F\begin{cases} .05 = 2.45 \\ .01 = 3.49 \end{cases}$ (37)

The F is larger than the .01 point, and it is clear that the students from high school *A* have varied greatly from year to year. The admitting college evidently cannot count upon receiving regularly groups of either good or poor students from this high school.

Dressel (**8**) has employed analysis of variance in attacking the question of the relationship between grades received in high school and later college success. In the 15 high schools studied, significant differences were found among the schools in grades assigned these students. The college grades received later by these students, grouped according to the high schools from which they came, varied much less than the wide differences in high school achievement might have led one to expect. Evidently, the marking standards employed by the high schools varied from school to school, and differed from those used by the colleges.

Jellinek (**23**) has used analysis of variance in several interesting physiological studies at the Worcester State Hospital. In one of these, the caloric output of 72 male schizophrenics was recorded under fasting conditions 12 (or less) times. The variations within and between individuals (called the *intra-* and *inter-*individual differences) were compared with the same measures obtained from normal controls. The inter-individual variance of the normals was significantly less than that of the patients. Jellinek writes: " . . . investigations in this hospital have shown that normal controls, although somewhat heterogeneous for most physiological variables, form a more compact group than schizophrenic patients" (p. 37).

The experimental design illustrated above has been employed most often in situations in which the groups compared were independent. If groups are *paired* initially person for person in some function, and later measured in the same or a related function; or if groups are matched initially (say for M and σ), analysis of vari-

ance may still be employed. Shen (**35**) has derived the formulas appropriate for use under these restricted conditions.

III. Multiple Classification of Variates

The method of analyzing the variance *between* and *within* groups can readily be extended to apply to more complex experimental designs. The chief difference lies in the larger number of groups needed and in the more complex relations among them. An example will serve to introduce certain new terms.

In one experiment (**11**), blood cholesterol readings (mg./100 cc.) were taken on 18 individuals during the months of April and May. The average reading for April was 144.91 and for May, 150.16. Is this rise significant? Data may be arranged as follows:

TABLE IV

Source of Variation	*DF*	*Sum Sqs.*	*Mean Sq.* (*V*)
Between months	1	248.0	248.0
Between individuals	17	4337.0	255.1
Interaction	17	1907.7	112.2
Total	35	6482.7	

$F = 248/112.2 = 2.21$ $F\begin{cases} .05 = 4.45 \\ .01 = 8.40 \end{cases}$ $F = 255.1/112.2 = 2.27$ $F\begin{cases} .05 = 2.30 \\ .01 = 3.30 \end{cases}$

It will be seen that instead of an analysis into *between* and *within* months, the within months variance has been separated into two parts: a new classification, "*between* individuals," and a term called "interaction." There are two determinations for each individual (one in April and one in May), and the new classification, "between the weighted means of individuals," enables us to find whether individuals differ significantly in blood cholesterol regardless of the month in which the reading is taken. Interaction is a measure of *experimental error*. This term gives a measure of the *per-month* fluctuations of the 18 subjects; it represents the variations arising from the fact that these subjects reacted differently on the two occasions. If each of our 18 *S*'s (subjects) had given the *same* reading in May as in April, *all* of the *V* would then be between individuals and there would be no interaction. Or if all *S*'s had returned identical readings in April and in May (each reading in May, however, being 10 points higher than in April), there would have been no interaction, *all* of the variation being then attributable to differences in the months (**18, 20, 39**).

Entering the table of *F*, first, with 1 and 17 *DF*'s (**37**), it is clear that the value 2.21 is *not* significant and that the blood

cholesterol readings did not really differ in April and May. It appears from the second F-test that individuals are almost—but not quite—significantly different in their blood cholesterol readings. An F of 2.30 might be expected to occur by chance only once in 20 trials, and the obtained F is 2.27.

Several aspects of this analysis may now be discussed at greater length. *First*, the total V has been analyzed into the V contributed by the *differences* in months, by the *heterogeneity* of the S's, and by the *interaction* of subjects on months. This last V is a measure of the failure of the S's to react in the *same way* to the experimental situation and hence is the most valid measure of the uncontrolled variation called experimental error. *Secondly*, it may be noted here that the F-test is essentially the same as the t-test (the critical ratio) when two groups only are to be compared. In fact, when two means are compared, $F=t^2$. In the problem above, the mean difference between months is 5.25 and the standard error of this mean is 3.53. Also, $t=5.25/3.53$ or 1.487, and for 17 DF (there are 18 differences) P is .156. Thus t^2 (*i.e.*, 1.487^2) is 2.21, which equals F. When several groups are to be compared, F is the better (more inclusive) test because it supplies an over-all test of significance. If F is *not* significant we need go no further. But if F is significant then the t-test should be applied to the separate comparisons, as F does not tell *which* differences are significant—but simply that some are. *Thirdly*, it should be remembered that when the measures or scores of two groups are pooled, equal σ's in the two groups are assumed, *i.e.*, the two groups are taken to be samples of equal variance drawn from a common population or from comparable populations. This condition, known as *homogeneity of variance*, is a prerequisite to analysis of variance. When means differ significantly, not only these parameters but the σ's within groups may differ, and the differences in σ may contribute to the mean differences. Tests for the homogeneity of variance (**21**) should be utilized when there are many groups and when doubt exists as to the comparability of their σ's. It is fortunate that the F-test is usually valid when the σ's of the groups to be compared are not equal (**26**), provided the discrepancy is not great.

The reader should note that had the analysis of variance in this problem been into *between* and *within* months, with 1 and 34 DF respectively, the experimental error would have been 183.7 (mean of 255.1 and 112.2) instead of 112.2. *All* subject variance—the within variance as well as the interaction variance—is now in-

cluded in this estimate of experimental error. It is clear that a design in which the experimental error is thus inflated decreases the precision of the result obtained.

The present methods of analysis of variance may often be usefully applied in experiments involving a number of practice trials.

Garrett (16), for instance, had *S*'s learn materials of different levels of difficulty under conditions of massed and spaced practice. One question raised by the experiment was whether sample (group) variance contributed more towards individual differences than did practice variance (*V* between means). The total variance in each set of practice trials was analyzed into the *V* between trials, the *V* within groups, and the interaction between *S*'s and practice trials. As was expected, both subject and practice variance contribute significantly in terms of experimental error (interaction). To test the relative strength of these two factors, practice variance was divided by sample variance and the resulting *F* evaluated in terms of the appropriate *DF*. In all cases but one, practice variance contributed more to total (*i.e.*, individual differences) than did subject variance, though not all differences were significant. The hypothesis is advanced that intensive practice is more potent than original nature (as measured by sample variance) in effecting differences in individual performance. But this conclusion is limited by the nature of the sample (college students) and by the character of the tasks.

A number of examples in which analysis of variance is an essential part of the experimental design may be found in the monograph by Jackson (21). One useful example (pp. 34–42) shows under what conditions several groups may be considered to be samples of a common population and combined into a single group.

IV. The Latin Square and the Greco-Latin Square

In some experiments the sequence or order of the experimental factors or situations in space or in time has a bearing upon the outcome. Agricultural is replete with such instances; and the spatial location of the plots in a field is often of great importance due to the possibility fertility gradients that may exist. Better results are expected from certain plots simply because of their location, and regardless of the treatment (*e.g.*, fertilizer).

In psychology, spatial position is often an important consideration, especially in experiments in psychophysics and in perception, and special methods have been devised to take account of this factor. Temporal order or sequence is also important and has received considerable attention in experimenta design. In fact, the rotation experiment in which groups are tested in regular sequence, or experimental factors introduced in some systematic order, was

devised with the avowed purpose of equalizing the time-order of experimental application. The experimental design known as the Latin Square provides a means of evaluating and eliminating the factors of spatial and of temporal sequence through analysis of variance methods. An example will illustrate the method.

In a recent experiment,* it was desired to determine whether the accuracy of recognition of colors differs markedly when the colors are presented to the dark adapted eye under various degrees of illumination. Colors selected were yellow, green, blue, and red. Four levels of illumination were chosen ranging from just no illumination through a graded series of intensities to a degree of illumination in which all subjects recognized correctly all the stimuli. Table V gives the experimental plan. The numerical entries in this table consist of the per cent of the experimental group who recognized each color correctly. Note that the table is so arranged that each color is presented *once* in the first, second, third, and fourth place at each of four levels of illumination; and *once* in first, second, third and fourth place in each series. This scheme produces a 4×4 table in which the fours columns and the four rows provide random arrangements of the colors. In consequence the variation in columns and rows represents a random variation which is independent of the influence of both order and of level

TABLE V

To Illustrate the Latin Square

The Four Colors, *R*, *G*, *Y*, and *B* are presented in such a way as to rotate out the Effects of Order and Position, and thus Equalize any Practice Effects.

Presentation	*Illumination Levels*			
Series	*1*	*2*	*3*	*4*
a	*R*	*B*	*Y*	*G*
b	*G*	*R*	*B*	*Y*
c	*B*	*Y*	*G*	*R*
d	*Y*	*G*	*R*	*B*

of illumination upon the recognition of the colors. This random variation in columns and rows will, however, decrease the average accuracy of performance. By segregating these influences, the effects of order and of illumination level are removed, and the accuracy with which the colors are recognized can be more precisely determined. Eliminating these sources of variation becomes, therefore, an effective means of error control. The analysis of variance takes the form shown in Table VI.

* Unpublished study by W. E. Vinacke.

TABLE VI

Source of Variation	*DF*	*Sum of Sqs.*	*Mean Square* (V)
Among rows	3	..	V (rows)
Among columns	3	..	V (columns)
Among colors	3	..	V (colors)
Residual error	6	..	V (error)
Total	15		

$$F = \frac{V \text{ (colors)}}{V \text{ (error)}}$$

It will be noted that the data have been analyzed into three sources of variation: intercolumns (illumination levels), inter-rows (order of presentation) and intercolors. Separated out also is the residual variance attributable to error. By means of the F-test, the variance of the accuracy of recognition of the different colors may be compared with the error variance, and the significance of the obtained differences determined. The variance due to rows (order) and the variance due to columns (illumination) can also be evaluated by the F-test. These factors are ordinarily of secondary interest in most experiments; but their size and their significance are often of value as indications of the effectiveness of our experimental controls.*

In another part of the experiment quoted above a study was made of the *form* of the stimulus instead of the color. An analysis similar to that in Table VI can be made of this situation, but it may be noted that such an experiment leaves unanswered the question as to the combined effect of color and form. To deal with this problem, we should have to use a Greco-Latin Square. In this experimental design, each color and each form are combined in such a way that only one combination of each sort occurs; and these combinations may then be treated as a Latin Square (**13, 41**).

V. Factorial Methods

Because they permit an experimenter to study simultaneously the action of two or more independent variables, the factorial methods of R. A. Fisher (**2, 18**) will prove especially valuable in many problems in experimental psychology. In factorial designs, each

* It should be noted that the use of percentages in an analysis of variance is not defensible, since the variance of a percentage itself ($\sigma^2 = pq$) and hence the variances of the columns and rows will not usually be equal. In order to circumvent this difficulty it would be well to transform the percentages into some other function whose variance is independent of the level of the functioning, *e.g.*, $\theta = 2 \sin^{-1}\sqrt{p}$; $\sigma_\theta = 1/\sqrt{N}$.

variable may occur at two or more levels or strengths, and the effects of the variables at each level (singly and in combination) may be investigated. Perhaps an example will provide the simplest introduction to this experimental design. Suppose that an experimenter (E) wishes to measure the relative strengths of thirst, hunger, and sex-deprivation in the white rat (**44**). The measure of each "drive" might be taken as the ***number of times*** a male white rat crossed an electrically charged "grid" after a given number of hours of deprivation from water, food, or sex-activity. For simplicity of design, let us suppose that each of our experimental variables occurs ar *two* levels, namely, *no* deprivation and x hours of deprivation. (More than two levels could, of course, be introduced.) We may set up our experiment as follows, the capital letters meaning presence of, and the small letters absence of the drives:

Variables	*Strengths or Levels*	
	Presence	Absence
Thirst	T	t
Hunger	H	h
Sex	S	s

There are 8 possible combinations of stimuli by which an animal may be affected: *THS*, *THs*, *Ths*, *ThS*, *tHS*, *tHs*, *thS*, *ths*. In the first situation, *i.e.*, *THS*, the animal is deprived of water, food, and sex; in the second, *THs*, of water and food, but not of sex, and so on. In the last experimental situation, *ths*, all three drives are at the zero level of deprevation, *i.e.*, at "normal." A total of 8 rats, one assigned to each experimental situation, is necessary to give the following form of analysis:

Source of Variation		*DF*	*Sum Sqs.*	*Mean Sq.*
Main Effects	T	1		
	H	1		
	S	1		
Interactions	TH	1		
	TS	1		
	HS	1		
	THS	1		
		7		

The letters T, H, and S, represent the "main effects" or the action

of the variables, taken separately, in causing the rats to cross the grid. *TH*, *TS*, *HS*, are the simple interactions of two factors, and *THS* is the one triple interaction. These interactions give the joint effects of two or of three factors in combination. Eight rats provide only one animal for each situation, and hence there is no measure of experimental error, as the outline above indicates. In order to provide a measure of the variability to be expected under the various conditions, the *E* would certainly "replicate" or repeat his experiment putting several rats through each variable-combination. Suppose that 8 rats were tested in ***each*** of the 8 experimental situations. The total group would then consist of 64 animals, the variance of whose "scores" (crossings) could be analyzed as follows:

Source of Variation	*DF*
T	1
H	1
S	1
TH	1
TS	1
HS	1
THS	1
Error	56
	63

One degree of freedom is assigned to each of the factors since each occurs at two strengths or levels. The *DF* of the interactions are simply the product of the *DF*'s assigned to the two (or three) interacting factors. There are 56 *DF*'s available for estimating the error variance, that is, the variability to be expected when several animals perform the same task. Note that if there had been 32 animals, providing 4 replications, there would have been 24 *DF*'s available for estimating error.

The significance of each factor may be found by dividing its *V* by the error *V* and evaluating the resulting *F* with due regard for the *DF*'s in the two *V*'s. The significance of the interactions is found in the same way.

It is worth considering further just what is meant by the significance of the main effects and of the interactions. Suppose that *H* is significant by the *F*-test. This would mean that hunger will cause a rat to cross the grid more often than lack of hunger; and

this is true when the other factors vary as to strength and combination. A significant interaction, say *HS*, means that hunger *and* sex-deprivation will cause a rat to attempt to cross the grid more often than will the normal rat ("normal" for these two conditions), whether he is thirsty or not.

The number of experimental variables and the levels at which they operate may be increased over those in our example. But it seems likely that 5 or 6 variables at 3 or 4 levels of each is about the limit of practical usefulness.

Crutchfield (**5**) (**6**) has carried out an experiment, using white rats, in which there were 5 factors each at 3 levels or strengths. The factors were concerned with various aspects of a string-pulling task, such as force of pull, length of string, training period, etc. Each experimental factor occurred at 3 levels. The outline of the experimental design was as follows:

Source of Variation		*DF*	*SS*	*MS(V)*
Single Factors	*A*	2		
	B	2		
	C	2		
	D	2		
	E	2		
Interactions,	*AB*	4		
Two factors	*AC*	4		
	AD	4		
	AE	4		
	BC	4		
	BD	4		
	BE	4		
	CD	4		
	CE	4		
	DE	4		
Complex Interactions; more than two factors (error)		192		
Total		242		

Since each factor occurs at 3 levels, each has 2 *DF*'s. Each simple interaction has 4 $DF(2\times2)$. Five factors each at 3 levels will provide 3^5 or 243 experimental situations; and hence 243 rats are needed in order to represent each experimental situation by *one* animal. The lack of replication would ordinarily be inadvisable, as it permits of no estimate of error. Where there are so many complex interactions, however, those of 3, 4, and 5 factors may be taken—as here—to give a fair measure of the differential reactions of the animals to the experimental situations, and hence provide an estimate of experimental error.

Crutchfield and Tolman (7) have summarized the advantages of factorial design over the conventional single-variable set-up. Their points will be repeated with some elaboration:

(1) *Factorial or multiple-variate design is more economical of time and of labor than is the single-variable plan, since several experimental factors may be studied simultaneously.* Furthermore, this design provides equal precision and demands no larger group of subjects than would be required if each experimental factor were studied alone. In Crutchfield's study 243 animals gave an appraisal of the 5 experimental factors, each at 3 levels, with as much accuracy as would have been obtained if 243 animals had been used to evaluate *each* factor singly, the other 4 factors being experimentally controlled. This means that 243×5 or 1215 animals would have been needed under the single-variable design to get a result obtained from 243 animals under factorial design.

(2) *Not only is there greater economy in factorial design, but such an experimental set-up is less artificial, and hence provides a more valid measure of the effects of each factor than does the single-variable method.* In the latter plan, results are valid only under the controlled conditions set up by the experiment. What effect a variable would have had under less restricted and more natural conditions, the experiment cannot tell. When several factors are controlled simultaneously or are held constant, conclusions relating to an experimental factor must be restricted to the given (and perhaps artificial) situation. In the illustration of p. 235, for instance, under single-variable method H would be studied while T and S were controlled; then T, while H and S were controlled, and so on. But these "drives" normally occur together and at various strengths. Factorial design provides (as here) a result valid over a wider range of conditions; single-variable design, on the other hand, provides only simple curves of performance at the different levels of the experimental factor.

(3) *Finally, factorial design supplies a means of studying the interactions or joint effects of the various factors.* These interactions may often be more important than the main effects. Also, the estimate of experimental error is usually more adequate in the factorial set-up, since the influence of possibly contributing factors is controlled through a segregation of the variance attributable to these causes.

Factorial or multiple-variable designs have application in a variety of psychological problems, such as the importance of various factors upon reaction time; of practice of different amounts and kinds upon transfer of training; of several variables at different strengths upon perception, memory, manual performance. In all of these, factorial methods, provide a broader base for generalization than can be obtained from single-variable methods. The main disadvantage of the factorial design is perhaps the complexity of the experimental situations when many variables at different strengths are to be investigated. Again, not as many levels of a

given factor can be tried out as is possible when only a single variable is studied at a time.

VI. The Applications of Analysis of Variance to Certain Special Problems

There are several applications of analysis of variance techniques to special problems which are important and useful enough to warrant special treatment.

1. *Reliability of mental tests.* The methods of variance analysis outlined in Section III are often useful in enabling an experimenter the better to evaluate a mental test or other measuring device. Suppose we have administered two forms (*A* and *B*) of the same test to 100 children in Grade V of an elementary school. Following the form of analysis given on p. 240 we may set up our data as shown in Table VII.

TABLE VII

Source of Variation	*DF*	*SS*	*MS (V)*
Between test forms	1	...	V_t
Between individuals	99	...	V_{ind}
Interaction (expl. error)	99	...	V_e
Total	199	...	

$F_1 = V_t/V_e$ $\qquad$ $F_2 = V_{ind}/V_e$

If F_1 is significant, then the performance of our subjects upon the two test forms is markedly different—due to practice or some other influence to be sought. If F_2 is significant, the test is a good measuring instrument, at least to the extent of separating individuals the one from the other. Should F_1 *not* be significant, practice and other effects have been negligible, and scores on the two forms are virtually equivalent. If F_2 is *not* significant, the test is clearly not differentiating the members of the group and hence is worthless.

In the study of a mental test, the analysis of variance may often be set up in another form more useful than that just presented. Thus the data in the problem outlined above might be arranged as follows:

TABLE VIII

Source of Variation	*DF*	*SS*	*MS (V)*
Between individuals	99	...	V_{ind}
Within individuals on the two forms	100	...	V_w
Total	199		

$F_2 = V_{ind}/V_w$

In this design the analysis of the total variance is into the V between individual subjects, for which there are 99 DF's, and within the scores of subjects on the two forms. Since each of the 100 subjects has two scores (one on Form A and one on Form B), the DF for "within" is 1 for each subject or 100 in all. Now if F_3 is not significant, the scores of individual subjects differ no more than do the scores made by the *same* individuals on the two forms of the test. Obviously, under these circumstances, the test is of no value. If F_3 is significant, the test is differentiating the members of the group.

From the V's in our table above, a statistic known as the *intraclass* correlation may be computed. Intraclass correlation differs from the usual "interclass" (product-moment) correlation. In interclass correlation, the variates X and Y fall into two columns and are distinct from each other: —r's between tests of intelligence and achievement, memory and reaction time are common examples. In intraclass correlation, there is no distinction between X and Y and the variables may be interchanged from one column to the other. The r's computed between the scores of twins in which either twin's score may be classified as X or as Y is an example of intraclass correlation. A method often used when variables are interchangeable is one in which each individual is classified first as X and then as Y, the r being computed from $2N$ cases. Other methods have also been devised (**19**).

In the design in Table VIII the intraclass correlation may be readily calculated from the formula,

$$r' = \frac{V_{\text{ind}} - V_w}{V_{\text{ind}} + V_w}.$$

When V_w is zero, *i.e.*, when the subjects' scores are identical on both forms of the test, $r' = 1.00$. A high positive r' indicates small variation within subjects' scores as compared to the variation between *different* subjects' scores on the two forms of the test. That is, each subject tends to score *alike* on the two forms, but individual subjects differ from each other. The intraclass r' is zero when the variation from subject to subject is no greater than is subject variability on the two forms. When this condition obtains, the same child is as unlike on Forms A and B as different children are unlike on the two forms.

The two attributes—consistency is subject score and differen-

tiation between subjects—are of course what one means by the reliability of a mental test. Hence, a high intraclass correlation gives more information regarding reliability than does a high interclass correlation. The latter (reliability coefficient or self-correlation), when high, tells us that the scores made by the *same* subjects on the two test forms are much alike *or* that they are different, but follow the same order. In the latter case, practice effects may cause the scores on the second test form to average 10 points higher, say, than on the first test. Consistency of score on the two occasions—in the sense of identity—is represented by an intraclass r' of 1.00, not by an interclass r of 1.00.

2. *Test for linearity of regression.* Analysis of variance may be readily utilized to test the linearity of regression in a correlation table. The test for linearity consists essentially in comparing the *V within arrays* (say within the *Y*-columns) with the *V around* the regression line, $\bar{y} = b_{yx} \cdot x$. If the regression line $\bar{y} = b_{yx} \cdot x$ passes through *all* of the column means, the difference between the sum of squares around the regression line and the sum of squares within columns is zero, and the regression is linear. The divergence of the sum of squares (within columns) from the sum of squares (regression lines) becomes greater as the regression departs more and more from linearity. When the departure is significant in terms of the *F*-test, the regression is nonlinear. Good examples and specimen calculations may be found in Lindquist (**26**) and Goulden (**18**).

3. *The standard error of an obtained score.* The usual formula for the standard error of an obtained score (sometimes called the standard error of measurement) is $\sigma_{1_\infty} = \sigma_1 \sqrt{1 - r_{1I}}$ where σ_1, the σ of the first form of the test, is assumed to be equal to σ_2, the σ of the second test form. The coefficient, r_{1I}, is the reliability coefficient of the test. When an analysis of variance of the form given in Table VII has been made of the test scores, the standard error of an obtained score may be found directly by extracting the square root of the *V* attributable to "error." This residual *V* is what is left when the variance due to repetition of the test and to characteristic subject differences is removed. It is, therefore, a valid measure of an individual's divergence from his theoretical "true" score.

4. *Item selection and validation.* Analysis of variance offers a promising approach to the problem of the selection and validation of test items, although it may be doubted whether the items of

most psychological tests and questionnaries are worth the time and effort necessary to validate them by this technique.

Baker (1) has sketched a plan for evaluating the items of a test by the method described in Section III. Let us suppose that the test items under study are to be validated against a criterion made up of school marks, achievement test scores, or other general measures. Total scores on the criterion measure may be separated into several (*e.g.*, 5) subgroups, described as *superior*, *above average*, *average*, *below average*, and *inferior*. Baker suggests that these groups include 10, 15, 50, 15, and 10 per cent, respectively, of the criterion group; and other percentages may, of course, be employed. The items to be selected are now administered to the entire criterion group, which is then divided into subgroups selected to be as nearly comparable as possible. The percentage of each subgroup answering a given item correctly is computed. The experimental set-up is as follows:

		Criterion Subgroups				
		1	*2*	*3*	*4*	*5*
Experimental	I	.85	.75	.50	.35	.10
subgroups	II	—	—	—	—	—
selected to be	III	—	—	—	—	—
comparable	IV	—	—	—	—	—

To illustrate for item No. 74, the percentages of experimental group I answering this item correctly are 85, 75, 50, 35, and 10 per cent when distributed into criterion subgroups. This means that in the subgroupings of experimental group I, 85 per cent of those classified as "superior" by the criterion scores answered this item correctly; 75 per cent of those classified as "above average" by the criterion scores answered the item correctly, and so on. The experimental design for the 5 criterion subgroups and the 4 experimental groupings is as follows:

Source of variation	*DF*	*SS*	*MS*(*V*)
Between criterion classes	4	—	*V*(cols)
Between exper. groups	3	—	*V*(rows)
Interaction (error)	12	—	*V*(error)
Total	19		

From the *F*-test one can determine whether an item varies significantly in percentage from one criterion subgroup to another. The variance introduced by the differences in experimental groups is eliminated by the design.

Lev (24) has used analysis of variance between and within groups (Section II) as a means of determining the weights to be assigned to each of three possible answers to a questionnaire item. Weights were assigned in accordance with the extent to which the three answers reflected truly different attitudes. Lev's method was to find whether the mean total score of those subjects answering *often*, *sometimes*, and *never* to a given item was significantly different in terms of the "within" variance. The

use of analysis of variance in situations like this is legitimate and interesting; but it is doubtful whether it is very practical. Experience of others with weighting systems (**25**), (**40**) indicates that a rule-of-thumb will usually give final scores which are as valid as those theoretically more precise.

5. *R. A. Fisher's "discriminant function."* The nature and purpose of Fisher's discriminant function can best be made clear by an illustration. Suppose that we have administered a battery of tests to two vocational groups, to groups of boys and girls, to abnormals and normals, or to any two contrasted groups. How best can we weight the tests of our battery so as to provide a maximum separation or discrimination between our two groups? The equation which yields composite scores under the condition of optimum weighting constitutes a "discriminant function." Such a function is essentially a regression equation from which—once we know a subject's score on our tests—we can estimate his "best" composite score. When we possess scores in some criterion measure against which our tests may be correlated, the discriminant function becomes an ordinary multiple regression equation.

Applications of the discriminant function to psychological problems—with some description of the mathematical derivation will be found in papers by Travers (**42**), Wilmott (**45**) and Lorge (27). Fisher's own account of this method (**12**) (**14**) will probably be difficult for psychologists who do not possess considerable mathematical training.

VII. Analysis of Covariance

The questions which analysis of *covariance* attempts to answer are quite similar to those with which analysis of variance is concerned. The main distinction between analysis of variance and analysis of covariance is the fact that in analysis of variance we have only ***one*** basic variable to analyze while in the analysis of covariance there are at least two such variables and in some cases many more than two. Thus analysis of variance may be regarded as *univariate* analysis, while analysis of covariance may be regarded as ***multivariate*** analysis.*

* This distinction between univariate and multivariate analysis must not be confused with simple and multiple classification for analysis of variance. (See page 235.) Another way of distinguishing between analysis of variance and covariance is to note that when one of the classification criteria in an analysis of variance is a continuous variable that bears a significant correlation to the basic variable and to one of the classification criteria, the procedure to be used is analysis of covariance.

It will be of interest to examine the situations in which it is desirable to modify the experimental design so as to include an analysis of covariance. As long as the experimenter can arrange the classification criteria at will and select at random individuals from each class, a simple analysis of variance is satisfactory. An analysis of covariance must be resorted to when a second variable, which is correlated with the basis variable under examination, shows a significant degree of association to one or more of the classification criteria. Thus, suppose we wish to determine whether there are any geographic or regional differences in the weights of children. The simplest approach would consist of dividing the total weight variance into two parts: *between regions* and *within regions* and then determine the *F* ratio for these two variances. Before concluding that weight is related to geographic location, we must determine whether age is a factor in this relationship. Age may be regarded as another classification criterion, and a two-way table can be drawn up in which weight is distributed according to the several subclasses of age in each geographic region. This gives rise to a two-way analysis in which the total weight variance is divided into three parts; between regions, between age groups and the residue which is independent of interregional and interage effects. The *F* ratio between the interregion and residual variance would then tell us whether or not there are regional differences in weight which are independent of the influence of age. This procedure is entirely valid if the age distribution is uniform in all the regions. But if age, like weight, also shows regional variation, and furthermore, if age is correlated with weight, the simple analysis of variance method does not apply and an analysis of covariance must be resorted to.

The simplest example of the analysis of covariance, and one especially useful in psychology, is the case in which the gains from initial to final performances in a given task are compared for two contrasted groups. Ordinarily, the gains in one group are compared with the gains in the second group by means of the critical ratio method. The method of simple analysis of variance could be utilized for this case. Such a procedure will answer the question as to whether or not the gain was related to the differential factor on the basis of which the two groups were selected or contrasted. But this answer will not be conclusive, because the difference in gain between the two groups may in part be due to whatever differences exist in the initial levels of the two groups.* In order

to be certain that the difference between the two groups in their respective gains is not influenced by the initial level of the groups, some way must be found to eliminate the influence of initial level from the gain.

The common practice utilized previous to the introduction of analysis of covariance was to use the regression equation between initial and final performances. By subtracting from the final performance of each individual the amount which was predictable from his initial performance, the residue independent of the initial level was obtained. These residuals could then be compared for the two contrasted groups, and the mean difference in residuals compared with the standard error of this mean difference.

This procedure of dividing the variance of one variable into two parts: (1) a part that is directly predictable from a second variable, and (2) a part that is independent of the second variable, is the essence of the method of analysis of covariance. It differs from the procedure just described in several essential points, but the underlying principle is the same.

The advantages that the analysis of covariance has over the older practice are: (1) it permits an over-all analysis when more than two groups are contrasted and (2) it permits a finer analysis in so far as it makes possible an examination of not only the total correlation and regression coefficients but also of the intergroup and intragroup coefficients. A simple example will illustrate the method.

In an experiment on the effect of convulsive electric shock on the memory of patients,** a series of tests consisting of paired word associates were taught before and after shock and the results compared with control data when no shock intervened.

The data are shown in Table IX.

Although an attempt was made to equate the tests for difficulty, this attempt did not prove successful. It is, therefore, important to remove the effect of initial difficulty from the final gain or loss in relearning. Table X gives the number of errors for tests under control and experimental conditions. Instead of predicting each relearning score from its initial learning score, it is

* It is well known that the correlation between gain and initial score is rarely zero. The exact relationship is given by $r_{x(x-y)} = (1 - kr_{xy})/\sqrt{(1 + k^2 - 2kr_{xy})}$ and this will vanish only when $r_{xy} = 1/k$, where $k = \sigma_y/\sigma_x$.

** These data were obtained from a series of mental patients treated with electric shock (**48**). Only selected data are presented to make the computation simpler.

TABLE IX

NUMBER OF ERRORS MADE IN LEARNING AND RELEARNING BY PATIENT P UNDER CONTROL AND UNDER EXPERIMENTAL CONDITIONS

Control			*Experimental*		
Series	*Learning* (*x*)	*Relearning* (*y*)	*Series*	*Learning* (*x*)	*Relearning* (*y*)
A	3	0	*G*	11	19
C	8	13	*K*	9	19
E	8	2	*M*	9	21
			L	39	31
Sums	19	15		68	90
Sums of squares	137	173		1804	2124
Sums of products	120			1778	

simpler to utilize a method which is in essence equivalent but which involves less computation (**26**). The only new type of sum introduced is the cross product. The sum of squares of the deviations are obtained in the usual way (**37**), and the correction factor which reduces the cross product to its deviation form is $[\sum(x)\sum(y)]/N$. These corrections are applied to reduce the sum of squares and products to deviation form. The results are shown in Table X.

TABLE X

SUMS OF SQUARES AND PRODUCTS FOR DATA IN TABLE IX. MEASURES ARE EXPRESSED AS DEVIATIONS FROM THEIR RESPECTIVE MEANS

Source	*Learning* (*x*)	*Relearning* (*y*)	*Cross Product* (*xy*)
Control	16.7	98.0	25.0
Experimental	648.0	99.0	248.0
Intra-Conditions Total	664.7	197.0	273.0
Inter-Conditions	195.1	525.0	320.0
Total	859.8	722.0	593.0

In order to analyze the data presented in Table X, it is necessary to recast it into two parts: the *inter*-condition variance, which gives the effect of the contrast in experimental conditions (control and shock); and the remaining or *intra*-condition variance, which gives the variability to be expected when all the learning and relearning are done under constant conditions. The within variance is obtained by squaring the deviation of each score, deviations in the two groups being taken from their respective means. The sum of squares for within conditions is shown in line three of Table X. The difference between the total and the within conditions sum of squares is equal to the between conditions sum of squares as shown in line four of Table X. These data have been rearranged

in Table XI to show the sums of squares and products for the two sources of variation (inter- and intra-conditions) divided by respective degrees of freedom.

TABLE XI

SUMS OF SQUARES AND PRODUCTS FOR DEVIATIONS FROM THE MEANS CORRESPONDING TO THE DATA IN TABLE X

Source	*DF*	*Learning Variance*	*Relearning Variance*	r_{xy}	b_{yx}	*Adjusted Sum of Squares for Relearning*	*DF*	*Adjusted Variance for Relearning*
Inter-Condition	1	195.1	525.0			295.4	1	295.4
Intra-Condition	5	132.9	39.4	.75	.41	84.6	4	21.1
Total	6			.92	.69	380.0	5	76.0
F (inter/intra)		1.5	13.3					14.0

In addition to the mean sums of squares and of products, Table XI also shows the correlation coefficients and regression coefficients for the intra-condition data and for the total data. The correlation is given by the equation:

$$r_{xy}=\sum(xy)/\sqrt{\sum(x)^2\sum(y)^2}$$

and the regression coefficient by the equation: $b_{yx}=\sum(xy)/\sum(x)^2$. (Data in Table X.)

There are three possible correlation coefficients and corresponding regression coefficients that could be calculated from Table X: (1) from the *inter*-condition data; (2) from the *intra*-condition data and (3) from the total. Since the inter-condition data reflect the influence of the experimental factor, it would be incorrect to utilize the coefficient from this source in estimating the intrinsic relationship between initial and final scores. The intra-condition correlation is suitable for this purpose, however, since it is free of any *systematic* effects of the experimental factor.* When the intra-condition regression coefficient was applied to the intra-condition and inter-condition sum of squares, the adjusted sum of squares (from which the influence of the initial level is eliminated) was 84.6 and 295.4 respectively. Similarly, when the intra-condition regression coefficient was applied to the total sum of squares the total adjusted sum of squares was found to be 380.0.** The differ-

* The total data cannot be used to determine the regression because these data include the effect of the inter-condition differences.

** The methods for performing this reduction ore given by Snedecor (37).

TABLE XII

ANALYSIS OF COVARIANCE FOR EFFECT OF SHOCK UPON MEMORY UNDER 4 EXPERIMENTAL CONDITIONS

Source	DF	Learning (x)		Relearning (y)		Cross Product	r_{xy}	b_{yx}	Unregressed or Adjusted		Regressed		Adjusted Variance
		Sum of Squares	Variance	Sum of Squares	Variance				Sum of Squares	DF	Sum of Squares	DF	
Inter-Condition	3	1293.6	431.2	1384.5	461.5	1132.1	.85	.88	393.7	2	990.8	1	196.9
Intra-Condition	15	1111.6	74.1	1244.5	83.0	698.7	.59	.63	805.3	14	439.2	1	57.5
Inter-Regression									36.4	1			36.4
Total	18	2405.2		2629.0		1830.8	.73	.76	1235.4	17	1393.6	1	72.7
$F=\frac{\text{inter-condition}}{\text{intra-condition}}$			5.8*		5.6*								3.4†

* F ratio significant.
† F ratio nonsignificant.

ence (342.0) between the unadjusted sum of squares for relearning (722.0) and the adjusted sum of squares (380.0) indicates the degree to which the initial level of learning controls the relearning variance. In assigning degrees of freedom to each source of variation, the usual procedure is followed of allowing one less than the number of independent measures in each category. It should be noted, however, that for the intra-condition sum of squares the difference between the adjusted and unadjusted (the regression effect) for relearning must have one degree of freedom assigned to it. Consequently, the 5 degrees of freedom that belong to the intra-condition sum of squares are distributed by giving 4 for the adjusted and 1 for the unadjusted sum of squares.

The first question to be investigated is: does the initial level of learning exert a sufficiently strong influence on the relearning level to require the adjusting provided by analysis of covariance? Let us examine the intra-condition variance. The unadjusted sum of squares of relearning from this source is 197.0 (Table X), and the adjusted sum of squares is 84.6 (Table XI), indicating that the adjustment has reduced the sum of squares by 112.4. The degrees of freedom corresponding to the unregressed* and regressed sum of squares are 4 and 1 respectively. Dividing the sums of squares by their respective degrees of freedom, we obtain the variances—21.1 and 112.4 for the unregressed and regressed variances. Dividing the latter by the former we obtain an F-ratio of 5.3, which for 4 and 1 degrees of freedom has a P value somewhat higher than .05. Consequently it is doubtful whether the initial level exerts a strong enough influence on the relearning level to affect the results. Usually one would drop the analysis at this point, but for the sake of finishing the illustration it will be carried out to its completion.

The second question to be answered is: does the removal of the influence of initial level produce any differences in the relationships between the two contrasted conditions for relearning? Let us first examine the unadjusted results.

When the unadjusted inter-condition variances are compared to the intra-condition variances, the F-ratios for the learning variances are not significant, but for relearning an almost significant

* That portion of the final variance which is predictable from the initial level is called the regressed variance, while the residue which is not predictable from the initial level is called the unregressed variance. The unregressed variance is identical with the adjusted variance.

ratio of 13 is obtained, indicating that the control mean probably differed significantly from the experimental mean in relearning. When the influence of initial learning is ruled out, the resulting *F*-ratio given by the adjusted variances is increased to 14, which is still in the doubtful zone of significance. Thus, the partialling out of the initial level does not materially change the significance of the comparisons.

So far we have dealt with a case involving simple analysis. Such cases can be treated as well by the ordinary methods of regression and group comparisons as by analysis of covariance, and the only innovation is the possibility of treating inter-group regression as well as intra-group regression.*

The power of the covariance method becomes clearer when there are more than two subclasses in the classification. The data previously presented in Table X were obtained from an experiment in which, in addition to the control and shock data, there were also inter-shock and post-shock data. These data are shown in Table XII. It will be noted that although the difference between the means of errors for the several experimental conditions was significant for the learning and relearning situations respectively, these differences disappear for relearning when the initial learning score (number of errors) is partialled out.

In Table XII each sum of squares for both sources of variation, the inter-condition and intra-condition, were adjusted by their own regression coefficients. The total sum of squares was also adjusted by its own regression coefficient. The adjusted or unregressed sum of squares as well as the regressed sums of squares are shown side by side for each source of variation. The *F*-ratio of regressed to unregressed variance is significant for the intra-condition source and for the total, indicating that the learning level exerts a significant influence on the relearning level.

One complicating element should be pointed out here regarding the problem of dealing with the three possible types of regression coefficients that could be utilized in adjusting the relearning means.** When the total of the adjusted sums of squares for the

* It will be readily noted that the inter-condition regression (or correlation) was previously regarded as "spurious correlation" due to heterogeneity of material. Yule (**47**). When two groups whose means are different are placed together in the same correlation chart, the resulting correlation will be spuriously high because of the heterogeneity that exists between the two groups.

** Although there are three possible correlation coefficients—inter, intra, and

two conditions was computed, it was found that it fell short by 36.4 of the amount computed by the total regression coefficient method. This amount of variation represents the variability that exists between the two regression coefficients and can be treated as another source of variation. Sometimes a question arises as to whether the intra-condition regression coefficient should be utilized to adjust the inter-condition and total, or whether each variance should be adjusted by its own regression coefficient (**22**). No general rule can be laid down, but it is advisable to utilize only one regression coefficient, because in that way only one degree of freedom is used up. Ordinarily, the regression coefficient for the total is to be preferred in such cases, since it is based on a larger number of degrees of freedom. However, if there is a significant inter-regression variance, it must be analyzed into its proper components.

Since in the present case the inter-regression variance was found to be only 36.4, which is even smaller than the intra-condition variance, it is possible to consider the inter-condition regression coefficient as not different from the intra-condition regression coefficient, and compute the residual variance accordingly. However, whether we utilize only one regression coefficient or both seems to make little difference in this problem, as the *F*-ratio still turns out to be nonsignificant.

Complexity is often introduced when more than one classification criterion is utilized. Thus, the data described above were obtained from several individuals. The total variance over all the conditions and all the individuals was analyzed not only for the inter-condition but also for the inter-individual effect.

A still further source of complication is that due to the introduction of additional variables. For example, besides the learning and relearning measures, there were also available recall* measures and recognition** measures on each of the tests. All of these additions are covered by easy extensions of the basic procedure of analysis of covariance, by means of which the variance of one variable over a series of subgroups can be analyzed after the influence of the correlated variables is eliminated.

---

total—only two of them are independent, since the total coefficient can be obtained from the combination of inter-and intra-condition coefficients.

* The recall is measured by the first trial in relearning.

** After the entire series was completed, a recognition test was given in multiple choice forms.

Examples of the analysis of covariance are even less frequent in psychological literature than are examples of analysis of variance. One of the earliest illustrations of the covariance technique was in connection with the scoring of tests of perseveration (**4**). This personality trait is usually measured by noting the decrease in speed of a given performance, *e.g.*, making horizontal tallies, when the directions are alternated between making horizontal and vertical tallies. The drop in speed indicates the perseveration of the preceding tendency. Usually the absolute difference between the control and experimentally altered speeds is taken as a measure of perseveration. It is apparent, however, that absolute difference will not be comparable from individual to individual because it depends on the normal level of speed in the function being measured. In some instances the ratio of this difference to the control speed is utilized in the hope that percentage differences in speed will be comparable from individual to individual. Rangachar (**29**) suggested that the best measure of perseveration is obtained by taking the difference between the observed experimental speed and the speed predicted from the control level. This is a direct application of the covariance technique.

Another instance in which the covariance technique might be useful is in connection with learning experiments involving control and experimental conditions when several measures of performance are available, such as trials, errors, and time. Instead of resorting to *a priori* weighting of these measures to obtain a total score, one of them (any one) can be selected as the index, and the observed measure compared with the measure predicted from the control levels in all three factors. The discrepancy between observed and predicted scores indicates the influence of the experimental factor.

Snedecor (**38**) has applied this technique to the problem of determining the homogeneity of teachers' marks. Struck by the difference between classes in their average attainment in mathematics, he set out to determine whether these differences reflect variation in mathematical ability or whether they reflect differences in teachers' judgments. It may be expected that the abler students will tend to have higher marks than the less able. Consequently, differences in marks between classes might well be laid to the influence of differences in initial ability. The two variables concerned were initial ability as determined by three preliminary tests, and the final mark in a mathematics course. Snedecor utilized the analysis of covariance technique to partial out the influence of initial level and discovered that differences between the classes still persist but are not nearly so marked as are the uncorrected differences. In short, the observed differences are due largely (but not entirely) to initial differences in ability and they are, in part at least, due also to differences in teachers' judgments.

Gaskill and Cox (**17**) have applied the method of analysis of covariance to studies of the influence of emotion on respiration. They investigated two variables: (1) the frequency of breathing (number of respirations per second), and (2) the number of atypical breaths per period under four experimental conditions. These were: (a) rest, (b) story read with abrupt and surprising ending, (c) crucial fear stimuli presented, and (d)

projection of various types of moving picture films. The problem may be stated as follows: "We want to know if the frequency of breathing has an influence upon the number of atypical breaths. If so, is the relationship the same under various emotional conditions?" It was found that the correlation between the two factors, frequency of breathing and number of atypical breaths, was only .003 for the entire set of experiments, but it varied considerably from experiment to experiment. The variance of the number of atypical breaths per period was analyzed after eliminating its relationship to the frequency of breaths. The adjusted variances showed a significant result for the inter-regression effect and for the inter-condition effect. Apparently each experiment produces a different effect on the number of atypical breaths even when frequency of breaths is kept constant.

The techniques of variance and covariance are applicable not only to experimentally controlled laboratory data such as those described previously, but they may also be applied under certain conditions to data over which the experimenter does not have complete control. One example of this is the twin study in which the experimental factors are predetermined and the task of the experimenter is observation rather than manipulation of experimental factors.

For example, Jackson (**21**) analyzed Wingfield's (**46**) data on resemblance of fraternal twins (unlike-sex pairs) in intelligence. In order to remove the differential effect of chronological age, the mental age scores were adjusted for chronological age. Results indicated that when the influence of chronological age is removed, the between-pair variance in mental age is no greater than the within-pair variance for pairs of twins of the same age. Therefore, the differences between twins are no greater than are the differences within twin pairs.

Several related or cognate methods, utilized frequently, must be clearly differentiated from analysis of variance and covariance. If scores on two forms of a test are available for each individual in two contrasted groups, two possible questions may be put to the data. The first concerns itself with the reliability of the test as a whole and its differential reliability in the two contrasted groups. The second concerns itself with discovering whether the gain from the first to the second test is different in the two contrasted groups. The first of these problems is a simple analysis of variance problem (p. 235) while the second calls for an analysis of covariance. Thus, the same data can be treated in different ways depending upon the purpose of the experiment.

A sharp distinction should be drawn between factorial analysis as carried on through the technique of analysis of variance and

covariance, and factor analysis as carried on through techniques which deal with correlational matrixes. Both of these procedures attempt to resolve a total variance into its components. The simplest distinction to be noted is that in analysis of variance, only one variable enters and is distributed according to various independent criteria; and the purpose of the analysis is to see how each independent classification criterion is related to the basic variable. In factor analysis, on the other hand, there are always several mutually dependent variables, and the purpose of the analysis is to determine underlying independent components that may reproduce the observed intercorrelations. Thus, the two methods are in a sense diametrically opposed, the first *beginning* with independent classification criteria, and the latter *ending* with independent factors. Since analysis of covariance reduces to analysis of variance in which several factors are held constant, the above statement holds equally true of covariance analysis. If we suspect that several factors may influence the outcome of a given experiment and can identify these factors beforehand, we can determine by means of analysis of variance whether the several factors actually influence the outcome. We can also determine whether a given combination of these factors has a differential effect (interaction) on the outcome. This situation is one which usually occurs in the laboratory where we can manipulate our factors to suit the experimental demands. When we do not know the factors that influence the outcome of an experiment, or when we know them but cannot control them experimentally, we can resort to factor analysis. Such situations are found in new research areas where little previous work has been done and there are many interrelated variables present. It is simpler in such cases to classify these variables into coherent groups by means of factor analysis, and in this way get some insight into the network of underlying relationships. Another way of distinguishing between the two techniques is to note that factor analysis deals with the differences between sets of tests for a given group of individuals, while variance analysis deals with differences in our tests between groups of individuals classified according to different sets of classification criteria.

We may close our discussion by summarizing those requirements which our data must satisfy before analysis of variance and covariance are suitable procedures. ***First,*** the individual measures or cases that fall into each category of the classification criteria

must be unselected. In order to determine whether any selection has taken place, the variances of each row and each column in the classification criteria must be approximately equal and whatever differences are found to exist should be attributable to random variation. *Secondly*, distributions obtained from several individuals, in which each individual's score enters more than once, should *not* be treated as independent data, since the scores of the same individuals are usually correlated. Variability within such data must be separated into two parts: inter-individual and intra-individual variance. *Thirdly*, the classification criteria must be independent. For example, if intelligence test data are classified by race and by region, the resulting table will not be amenable directly to the analysis of variance techniques because there is definite association between the proportion of Negroes and regional distribution of the population. An estimation of error derived from such analysis will be swelled by this correlation and any conclusion based on the analysis, although erring on the conservative side will not lead to conclusive results. *Fourthly*, the basic variable under examination must have a normal distribution in the population from which the sample is drawn. *Fifthly*, for analysis of covariance, the correlations must be linear.

The extent to which deviation from these requirements affects the results of the analysis have been examined by Pearson (**28**), who concluded that even when some of the above requirements are not rigorously adhered to, the results of the analysis may still be meaningful. The degree of deviation from the prescribed requirements which would nullify the conclusion based on the analysis is still to be determined.

BIBLIOGRAPHY

1. Baker, K. H. Item validity by the analysis of variance: an outline of method. *Psychol. Rec.*, 1939, **3**, 242–248.
2. Baxter, B. The application of factorial design to a psychological problem. *Psychol. Rev.*, 1940, **47**, 494–500.
3. Baxter, B. Problems in the planning of psychological experiments. *Amer. J. Psychol.*, 1941, **54**, 270–280.
4. Cattell, R. B. Temperament tests. II Tests. *Brit. J. Psychol.*, 1933–34, **24**, 20–49.
5. Crutchfield, R. S. Efficient factorial design and analysis of variance illustrated in psychological experimentation. *J. Psychol.*, 1938, **5**, 339–346.
6. Crutchfield, R. S. The determiners of energy expenditure in string-pulling by the rat. *J. Psychol.*, 1939, **7**, 163–178.
7. Crutchfield, R. S., & Tolman E. C. Multiple-variable design for experiments involving interaction of behavior. *Psychol Rev.*, 1940, **47**, 38–42.

8. DRESSEL, P. The effect of high school on college grades. *J. educ. Psychol.*, 1939, **30**, 611–616.
9. DUNLAP, J. W. Applications of analysis of variance to educational problems. *J. educ. Res.*, 1940, **33**, 434–442.
10. ENGELHART, M. D. The analysis of variance and covariance techniques in relation to the conventional formulas for the standard error of a difference. *Psychometrika*, 1941, **6**, 221–233.
11. FERTIG, J. W. The use of interaction in the removal of correlated variation. *Biom. Bull.*, 1936, **1**, 1–14.
12. FISHER, R. A. The use of multiple measurements in taxonomic problems. *Ann. of Eugenics*, 1936, **7**, 179–188.
13. FISHER, R. A. The design of experiments. London: Oliver & Boyd, 2nd edition, 1937.
14. FISHER, R. A. The statistical utilization of multiple measurements. *Ann. of Eugenics*, 1938, **8**, 376–386.
15. FISHER, R. A. Statistical methods for research workers. London: Oliver & Boyd, 7th edition, 1938.
16. GARRETT, H. E. Variability in learning under massed and spaced practice. *J. exp. Psychol.*, 1940, **26**, 547–567.
17. GASKILL, H. V., & COX, G. M. I. Respiration; use of analysis of variance and covariance in psychological data. *J. gen. Psychol.*, 1937, **16**, 21–38.
18. GOULDEN, C. H. Methods of statistical analysis. New York: Wiley, 1939.
19. GUILFORD, J. P. Psychometric methods. New York: McGraw-Hill, 1936.
20. HACKMAN, R. B. An experimental study of variability in ocular latency. *J. exp. Psychol.*, 1940, **27**, 546–558.
21. JACKSON, R. W. B. Application of the analysis of variance and convariance methods to educational problems. *Dept. Educ. Research. Univ. Toronto Bull.*, 1940, **11**, 67–74.
22. JACKSON, R. W. B. Some difficulties in the application of the analysis of covariance method to educational problems. *J. educ. Psychol.*, 1941, **32**, 414–422.
23. JELLINEK, E. M. Measurements of the consistency of fasting oxygen consumption rates in schizophrenic patients and normal controls. *Biom. Bull.*, 1936, **1**, 15–43.
24. LEV, J. Evaluation of test items by the method of analysis of variance. *J. educ. Psychol*, 1938, **29**, 623–630.
25. LIKERT, R. Technique for the measurement of attitudes. *Arch. Psychol.*, 1932, **22**, No. 140, 55 pp.
26. LINDQUIST, E. F. Statistical analysis in educational research. Boston: Houghton Mifflin, 1940.
27. LORGE, I. Two-group comparisons by multivariate analysis. *Amer. Educ. Res. Assoc., Official Report*, 1940, 117–121.
28. PEARSON, E. S. The analysis of variance in cases of non-normal variation. *Biometrika*, 1931, **23**, 114–133.
29. RANGACHAR, C. Differences in perseveration among Jewish and English boys. *Brit. J. educ. Psychol.*, 1932, **2**, 199–211.
30. RIDER, P. R. An introduction to modern statistical methods. New York: Wiley, 1939.

31. RIETZ, W. Statistical techniques for the study of institutional differences. *J. exp. Educ.*, 1934, **3**, 11–24.
32. ROFF, M., & ROFF, L. An analysis of the variance of conflict behavior in preschool children. *Child Develpm.*, 1940, **11**, 43–60.
33. RUBIN-RABSON, G. The influence of analytical pre-study in memorizing piano music. *Arch. Psychol.*, 1937, **31**, No. 220. 53 pp.
34. SCHRADER, W. A. B. Analysis of variance applied to liberalism scores. *J. exp. Educ.*, 1940, **8**, 267–270.
35. SHEN, E. Experimental design and statistical treatment in educational research. *J. exp. Educ.*, 1940, **8**, 346–353.
36. SHUTTLEWORTH, F. K. The nature and nurture problem, Part II. *J. educ. Psychol.*, 1935, **26**, 655–681.
37. SNEDECOR, G. W. Calculation and interpretation of the analysis of variance and covariance. Ames, Iowa: Collegiate Press, 1934.
38. SNEDECOR, G. W. Analysis of covariance of statistically controlled grades. *J. Amer. statist. Ass.*, 1935, **30**, 263–268.
39. SNEDECOR, G. W. Statistical methods. Ames, Iowa: Collegiate Press, 2nd edition, 1938.
40. STRONG, E. K. Interest maturity. *Personnel Journal*, 1933, **12**, 77–90.
41. THOMSON, G. H. The use of the Latin square in designing educational experiments. *Brit. J. educ. Psychol.*, 1941, **11**, 135–137.
42. TRAVERS, R. M. W. The use of the discriminant function in the treatment of psychological group differences. *Psychometrika*, 1939, **4**, 25–32.
43. WALKER, H. M. Degrees of freedom. *J. educ. Psychol.*, 1940, **31**, 253–269.
44. WARDEN, C. J. Animal motivation; experimental studies on the albino rat. New York: Columbia Univ. Press, 1931.
45. WILLMOTT, J. N. High school boys electing industrial arts; a study of certain factors differentiating the industrial arts group from the group not electing industrial arts. New York: Teach. Coll. *Cont. Educ.*, 1941, No. 836.
46. WINGFIELD, A. H. Twins and orphans; the inheritance of intelligence. London: J. M. Dent & Sons, 1928.
47. YULE, G. U., & KENDALL. M. G. An introduction to the theory of statistics. London: Charles Griffin & Co., 1937.
48. ZUBIN, J., & BARRERA, S. E. Effect of electric convulsive therapy on memory. *Proc. Soc. Exp. Biol. N. Y.*, 1941, **48**, 596–597.

PSYCHOLOGY AND THE WAR

Edited by

STEUART HENDERSON BRITT

CONTENTS

PERSONNEL RESEARCH IN THE ARMY. III. SOME FACTORS AFFECTING RESEARCH IN THE ARMY, by *Staff, Personnel Research Section, Classification and Replacement Branch, The Adjutant General's Office* 271

THE SPECIAL TRAINING UNITS OF THE ARMY, by *Morton A. Seidenfeld* 279

THE MILITARY USE OF THE RAIL-WALKING TEST AS AN INDEX OF LOCOMOTOR COORDINATION, by *S. Roy Heath, Jr.* 282

A COURSE IN ARMY PERSONNEL TECHNIQUES, by *Calvin S. Hall* 285

THE COMMITTEE ON FOOD HABITS, by *Margaret Mead* 290

ATTITUDES OF YOUTHS OF HIGH SCHOOL AGE TOWARD THE WAR, by *Mandel Sherman* 294

PSYCHOLOGY AND THE WAR: NOTES 300

WOMEN PSYCHOLOGISTS AND THE ARMED SERVICES
PSYCHOLOGISTS IN WAR WORK
REGISTRATION WITH OFFICE OF PSYCHOLOGICAL PERSONNEL
COMMITTEE ON WAR SERVICES TO CHILDREN

PERSONNEL RESEARCH IN THE ARMY
III. SOME FACTORS AFFECTING RESEARCH IN THE ARMY

BY STAFF, PERSONNEL RESEARCH SECTION, CLASSIFICATION AND REPLACEMENT BRANCH*

The Adjutant General's Office

Personnel research in the Army, like that in any specialized field, has its own special characteristics. Military organizations and conditions differ markedly from civilian groups and situations; consequently the methods used in any military activity are subject to factors affecting to a marked degree the work which is done. These factors operate in two ways: by creating possibilities generally not found in personnel research and by imposing limitations which must be taken into consideration in every undertaking. The purpose of this article is to afford some insight into the framework within which studues are conducted, including both possibilities and limitations. This furnishes a background for understanding better the reports of specific studies carried on by the Personnel Research Section to be reported subsequently.

In spite of the great advances made in military personnel practices during the last war, it can hardly be said that any well-defined body of content or procedures existed when the present war began. Little scientific work in Army testing was done between the two wars, whereas other aspects of the field of applied psychology advanced conspicuously. This made it impossible to begin where Army personnel work had stopped in 1918. The progress made during the last war and advances made in related fields since were used as the basis for beginning to study the new personnel problems of this war. Leaning heavily on recent developments in psychology, the Personnel Research Section has adapted those which are applicable to the new situation. A new group of workers has had to become familiar with military practices, for the Army

* The title of the Branch has been changed during the past month from Classification and Enlisted Replacement Branch since it has been given responsibility for the replacement of officers as well as of enlisted men.

did not employ any psychologists until the enlargement of our armed forces began in 1940. The fact that testing and classification work are now combined in The Adjutant General's Office has made possible a more integrated approach to Army personnel problems, and a new phase of applied psychology is being intensively developed.

The Personnel Research Section came into being because it was essential to the winning of the war. It was not set up to carry on research for its own sake. Rather, research is carried on as a service to the War Department, and every problem that is studied is born of a real and urgent need. Only those problems directly related to the war effort can be investigated. This vitalizes the entire program. Freedom and human lives, not theories, are at stake. The reality of the problems, the necessity for action, the consequences that are involved, the magnitude of the enterprise, all give quality to the work that cannot be dissociated from the results obtained. In the following sections are discussed some of the problems affecting personnel research in the Army.

Size of the Army. The most obvious influencing factor is that of sheer size. The Army is engaged in the largest personnel program in the history of the country, perhaps of the world, and size brings its own peculiar type of possibilities and difficulties. A partial idea of the Army's size, and its attendant problems, may be gained from imagining the situation that would result if all the colleges and universities in the country were brought under one centralized administrative head. Even then, the resulting organization would be less than one-third as large as the Army will be at its maximum. In addition, the personnel policies in a university affect the lives of students to only a fraction of the extent that the Army's personnel policies affect the lives of soldiers.

To speak of the sample on which research is being conducted in the Army is hardly accurate. It would be more proper to refer to it as a population, for with few exceptions data are being collected on all males between the ages of 18 and 38 eligible for military service. The number will be far greater than in the last war. In 1917–18 the Army attained a size of only about three million in comparison to the seven or eight million man Army now being built. Furthermore, data will be available on all the men that are being inducted now. In contrast, during the last war data were available on less than half the men, since the psychological measures were not ready for administration until long after our country

had entered the conflict. Number, in and of itself, means little but it has certain advantages. It makes possible an accurate check on the reliability of all the measures which are devised and permits the comparison of men from varying geographical areas to a degree never before possible. It has also necessitated the development of personnel procedures adapted to large groups which may serve many useful purposes in the future. A by-product of the size of the Army is the fact that at any one time large samples of almost any description are available. For example, there are always a large number of men being inducted, another group finishing basic training, another beginning officer training. Data with reference to a particular measure can thus be collected with great speed, and the results of any research can be put into practice with the same rapidity.

With the formation of the Women's Army Auxiliary Corps, data are also being collected on women. Thus an entirely new area of research is being opened. The number of women which it is planned to induct into the Auxiliary Corps indicates that the size of the sample should be adequate for obtaining significant results.

Diversity of Problems. The highly specialized nature of the present war has necessitated research on a variety of problems perhaps never before equalled in one organization. In this respect the size of the Army has been a definite spur to the construction of precise measuring instruments. Correct placement has become an urgent necessity in order that an organization the size of the Army can function at its necessary efficiency.

Mention of a few of the problems will give some indication of their diversity and scope. The fundamental job was the construction of tests to give an index of ability to learn, for general classification of all men inducted into the Army. The Army General Classification Test now serves that purpose. For those men who neither speak nor write English or who are illiterate, a Visual Classification Test was devised.

The many specialized jobs necessary in the Army have led to the construction of tests in a variety of different fields, such as those for radiotelegraph operators and truck drivers. In order that men with trade backgrounds can be correctly selected and placed, and given the training they need, trade knowledge and aptitude tests have been devised. For various occupations, educational achievement tests were necessary. The selection of officer candidates has brought about research on leadership and a large number

of tests have been built for the selection of warrant officers. Interrelated with the construction of tests is the work on occupational analysis, carried on as part of the duties of the Classification and Replacement Branch. Military occupations in all arms and services are analyzed, and data are calculated and kept on the occurrences of and requirements for all military specialists.

Time Available. The pressure of time pervades every problem and solution. This has often been called a war against time as much as against the enemy. Time saved—or lost—is measurable in terms of human life and national security, rather than as educational or industrial efficiency. All training programs are tightly planned. Every moment is given over to teaching enlisted men and officers knowledge and skills which will be of use in combat. There is no time allotted to research as such, and each interruption for the purpose of collecting data breaks into the schedule far more seriously than it would in education or industry. Taking a morning or an afternoon for the administration of a test dislocates the training program. It may mean that the soldiers in that group miss instruction of considerable importance. Because of this, the administration of tests not called for in the regular routine generally takes place in the evening after the day's duties are done.

Chemical and physical research differ from personnel experimentation and research in that they can proceed largely without disrupting training schedules. In personnel studies large amounts of soldiers' time are necessarily involved, and time is of first importance. There are situations in which even a few hours taken from a training program to collect data not immediately useful to the training officer may appear to be time ill-spent when viewed only in terms of the situation at hand. The proper authorities must be convinced that the present sacrifice will pay dividends in the near future. In research programs carried on over a long period of time, such difficulties are multiplied. However, these are not insurmountable. Data for the study of any problem can always be collected even though training programs are interrupted. This is, however, a factor that must be taken into consideration in the gathering of data.

Army Organization. For purposes of research, military organization has many advantages. The uniformity which exists in Army installations makes possible the collection of comparable data from many sources. Once the necessity for obtaining certain

data is apparent, a directive sent to the field is assured of compliance. One can be certain that the same records are kept on a soldier in Maine as on one in Montana, and that information on both is equally accessible.

A complete record of the Army experience of every soldier is kept on the Qualification Card. All test scores, course grades, and assignments are recorded. In addition, considerable material on a man's civilian experience is collected when he is inducted. A wealth of data is thus available for every subject in a study, making possible the investigation of relationships seldom possible in other situations.

Selecting and Maintaining Samples. The system of quotas that exists in the Army has a direct effect on the quality of the samples in studies. Based on the needs of the various arms and services, quotas are set up over certain periods. It may be that because of imminent military operations a large number of men will be needed in the Engineer Corps. When that is the case, the quota is filled from the men who are then being inducted. Naturally, because of this system, some misplacements result. This may mean only that a man is placed in the field of his second rather than his first choice, and reassignment of men incorrectly placed is made during the first few months of training. The assignment of men where needed is, however, an understandable military necessity. Within the number of men referred to above who are classified for training in the Engineer Corps, certain other quotas are set; that is, a specified number must be trained as officers and others must be trained as specialists (in demolition, bridge building, etc.) to provide the skills needed in the projected operations. If the quotas for the Engineer Corps are high over a period of time, the average caliber of the men assigned to the engineers will probably drop; and, when men already in the Corps are assigned to specialist training, it is likely that as quotas are increased more men of low ability will have to be selected for that training. Thus, the populations of two groups in the same type of training may be quite different.

As can be gathered from the above, there is frequently a conflict between the quality of men wanted for a certain type of training or position and the men who are available. In such cases a sacrifice of one or the other is made. Either men of lower aptitude are accepted and given all the training possible in the available time, or fewer men are taken than originally requested and no change made in the training standards. In general, only men in

grades I, II, and III on the Army General Classification Test* are selected for specialist training, and men in grades IV and V supply the bulk of the basic soldiers. There are exceptions to this, however. If the civilian experience of a man is one that is needed in the Army, he is placed in that job regardless of his Army grade. If a man has made a high score on an aptitude test, his chances of being selected for that type of training are high regardless of his grade in the General Classification Test.

The effect of the quota system on populations in studies is often considerable. Over a period of time, classes in a particular subject vary widely with regard to certain measures. For example, in one training school the mean scores of five classes on a code learning test varied from 38 to 52, with no group including less than 100 cases, the standard deviations for all groups varying within four points, from 18 to 22. In another instance, a group of infantrymen being trained as radio operators had attended school an average of 10.65 years, with a standard deviation of 2.65 years; a group of airmen at another camp, also being taught code, had attended school an average of 12.54 years, with a standard of only 1.68 years. Examples such as these could be multiplied to illustrate the difficulties of procuring samples which are unselected and the caution which should be used in comparing results from one group to another. Unselected samples are obtainable at Reception Centers for, with few exceptions, all men going into the Army pass through these. For research on general measures, the men going through the Reception Centers are used. The problem, however, of securing an unselected sample for validation of measures in relation to training is frequently a difficult one.

Determining Validity Criteria. Perhaps the most acute difficulty of all is that of securing adequate validity criteria on the tests which are constructed. This problem is not unique nor limited to the Army. The ultimate criterion for all military tests is the behavior of soldiers in the jobs for which they are trained, namely, success in combat. Validation data of this sort have, to date, been impossible to collect, as the Section has been limited in its work to the continental United States. Even if arrangements were made for the evaluation of soldiers in combat with respect to their special jobs, the problems of collecting such data would be enormous and

* The Army General Classification Test is given to all men entering the Service. On the basis of score, each man is given an Army Grade. These range from I to V, I being high.

the control of conditions exceedingly difficult. However, a project of that sort is not out of the question, and it is hoped that work of that nature may be undertaken in the near future.

In the absence of final validity criteria, others must be used. In some instances these criteria are easy to secure and relatively satisfactory. For radiotelegraph operators, the number of words per minute which a man can send or receive in code is definite, readily obtainable, and highly valid in terms of instruction in a communications course. Whether or not the performance of a soldier in sending and receiving code would be altered much by combat conditions cannot at present be determined. In radiotelegraph courses, much of the instruction is concerned with non-code material. Here adequate validity criteria are more difficult to secure, and grades given by instructors are used. As is true with course grades in any situation, their validity is often questionable.

Another common practice in Army instruction is to pass all, or nearly all, the soldiers in some types of training. Although it is recognized that a wide variation in content learned or skills acquired may exist in a particular class after a course, practically no men are failed because it is felt that all have mastered certain fundamental concepts or techniques enabling them to meet situations which will arise. Where this is the case, each man is given a passing grade or better in the course. By thus cutting off all the cases at the lower end of a distribution, the validity coefficients of a particular measure are substantially reduced.

Field Personnel. Securing competent personnel for work in the field was, for a time, a difficult problem. Large numbers of trained men are needed, and not enough were being inducted to carry on the work in the many Army installations. Augmenting the problem was the shortage of professionals in the Washington office, which made close supervision a virtual impossibility. There were, of course, many highly trained officers and enlisted men in charge of personnel work in the field, but it was also necessary to recruit additional men with less background for the work than was desired. However, The Adjutant General's Office has set up schools for officers, officer candidates, and enlisted men; and many thousands have already been given intensive training in personnel and classification work. Here again, the organization of the Army has been a distinct asset, for it has been possible to control the training which these men have received and to fit them for the positions which they will occupy. In addition to the schools, many of the

officers in the field have trained their own staffs. The quality of the field personnel is generally of high caliber, and they are doing effective and efficient work.

Summary

Procedures in Army personnel research are determined by several factors, unique to the Army, which both open possibilities and impose limitations: the size of the Army, the wide variety of problems on which research is being carried forward, the pressure of time, Army organization, selecting and maintaining samples, determining validity criteria, and the building up and training of a competent field staff. A knowledge of these factors and how they operate is essential to an understanding of how personnel work is conducted in the Army and also of the results that are obtained.

THE SPECIAL TRAINING UNITS OF THE ARMY

BY MORTON A. SEIDENFELD
Lt. Colonel, Adjutant General's Department

Generally when one thinks about training in the Army, there are visualized columns of men in uniform, drilling with rifles or smartly marching down the field with heads held high and feet sharply stamping out cadence with staccato clarity. Such a mental picture is oftentimes correct; yet to be completely accurate it should include an occasional man who just cannot keep in step, who does not know "right" from "left," or one who marches with a decided limp.

Within the military services there are a sizeable number of men who are moderately physically handicapped, some who are on the dull side mentally, and still others whose learning is retarded due to lack of acquaintance with the English language. Such men, however, are capable of giving a great deal of useful military service provided they are trained to execute the duties assigned to them. This means that each man in these categories must be given some specialized form of training prior to, or concurrent with, their regular military instruction. Obviously such training requires more time than normally allocated to basic military instruction within the Army. This additional time for training and the special program of instruction is provided by the establishment of an educational program designed specifically for Special Training Units. Men assigned to such units, in addition and prior to regular training, take from eight to thirteen weeks special training during which half-time is devoted to academic instruction in the fundamental subjects of reading, writing, and arithmetic, and the remaining half to instruction in the basic military subjects.

In the case of men who are assigned to Special Training Units as a result of physical deficiencies only, the program of training is based upon the concept of teaching the man how to overcome his handicaps. Many men are able to perform "general service" in spite of physical limitations; others, however, are able to perform only a portion of the duties of a soldier, and are assigned to "limited

service." If a man cannot accomplish his duties at first, he is given necessary stamina-building exercises so that he can successfully meet such situations as may ordinarily be encountered in the assignments for which he is considered physically qualified. Such a program is in actuality a buffer, affording the man adjusted to his civilian environment an opportunity to make the transition to military life without breaking his morale.

Similarly, but using techniques appropriate to the problems involved, men who are in the slow learner group (mentally dull) are provided not only with the opportunity to be trained under conditions that are pitched at their capacity to acquire knowledge, but they are also given instruction by personnel who are especially qualified to teach men in this category.

Those who are in the foreign-language group, who cannot read or write English to a degree which will satisfy Army requirements, are likewise taught the fundamentals of the language within the Special Training Unit.

To accomplish this program, the War Department has developed textbooks and instructional materials, including teachers' guides and visual aids necessary to assist instructors. The Section of the Adjutant General's Office concerned with adult elementary education in the military services is carrying on research to develop effective and rapid techniques. With training time relatively brief, all methods are pointed toward establishing maximum efficiency in the learning process. The materials developed place particular emphasis on motivating the learner toward an eagerness to learn all he can as rapidly as possible. His academic learning is designed to tie in well with his military training so there will be no lost motion in carrying him from these subjects into the task of learning to be a soldier.

The instructional staff at the various military installations having Special Training Units is composed of both officers and enlisted men. A very large proportion of this personnel has a background of training in the field of Education augmented by teaching experience. Unfortunately, the problem is so large in scope that it has been difficult to secure a sufficient number of teachers with special training and experience. Many qualified teachers are actively engaged in carrying on other instructional activities and cannot be assigned to this work. Accordingly a course of instruction on Methods of Teaching in Special Training Units is now being

prepared for the field by The Army Institute, and a school is contemplated for teacher training of instructors in these units.

The staff of officers and professional civilian personnel dealing with Special Training Units within the War Department has produced a great deal of original instructional material in the form of textbooks, film strips, periodicals, and current event leaflets, as well as aids and guides for instructors. These materials are characterized largely by their freedom from rigid and static content. Every effort has been made to maintain an eclectic attitude so that teacher and student alike will be provided with ample opportunity to teach and learn in a manner most effective to their individual backgrounds. Much emphasis has been placed upon visual aids as a valuable auxiliary in training; auditory, visual, and kinaesthetic sensory mechanisms are used to the fullest extent to augment each other and to assure the learner of means that will be adapted to his particular learning pattern. Such an approach is essential in view of the limitations on the amount of time available for training. The evidence indicates that with present techniques about 95 per cent of all the men sent to special training successfully complete their instruction and go on to regular training.

Through the special training program, many thousands of men have already been made more useful in the Army; in fact, many men would have had to be discharged had it not been possible to bring them up to an acceptable training level through the efforts of officers and men carrying on this program in the field.

THE MILITARY USE OF THE RAIL-WALKING TEST AS AN INDEX OF LOCOMOTOR COORDINATION

BY S. ROY HEATH, JR.

First Lieutenant, U.S.A., Personnel Consultant, Field Artillery Replacement Training Center, Fort Bragg, N. C.

Good muscular coordination is readily recognized among military men as a prerequisite for soldiers selected for tactical duty. The success with which the Rail-walking Test (1) scores discriminate between groups approved for tactical duty and those not approved reenforces the validity of the test as an index of locomotor coordination. If a soldier fails to react smoothly and quickly in his physical movements, his presence as a member of a tactical team will be costly for both manpower and materiel. The poorly coordinated are better used for duty in the non-tactical sphere where less precision and speed of movement are required. Intelligence and emotional stability are also differentiating factors but motor coordination remains one of the most critical.

The Rail-walking Test consists of walking three wooden rails 9 feet by 4 inches, 9 feet by 2 inches, and 6 feet by 1 inch, of sufficient height to prevent contact of foot with floor. Each rail is graded in units of feet and is walked barefooted heel to toe by the subject three times. The raw score is the total number of feet walked without falling off. Because of the difference in degree of difficulty of each rail the raw score for the three rails is weighted in a 1–2–4 ratio. Thus, the total possible score is (1×27) plus (2×27) plus (4×18), or 153 points. The test was standardized on 1,013 troops selected from a non-special unit which could be regarded as an average cross section of the white male population of Army age. The mean Rail-walking score for this group was 130.73 points with a standard deviation of 19.63 points.

Seventy-three soldiers selected at random from a group (not included in the normal standardization) who after initial training were classified by the Army as "able to meet the minimum standards for tactical units" were tested on the Rail-walking Test. The mean score for this group was 134.41 points with a standard deviation of 13.70 points. No score was below 94 points. Next 94 soldiers, selected at random, were tested from a group which was composed

entirely of men found unsuitable by the Army for duty in tactical situations. The mean Rail-walking score was 86.06 points with a standard deviation of 32.80 points. Fifty-three cases, or more than 50 per cent, fell below 94 points in score.

In another situation the Rail-walking Test was administered to 48 soldiers who were referred as problems of locomotor coordination to the Special Training Unit at the Fort Bragg Field Artillery Replacement Training Center. The mean score for this group was 68.98 points with a standard deviation of 35.60 points. Thirty-nine, or over 80 per cent, of the cases fell below 94 points.

There are, however, several reservations which must be made in the gross use of the Rail-walking Test as an index of locomotor coordination. During training experience with problem soldiers in the Special Training Unit the author has observed that all persons with coordination problems do not necessarily have a low Rail-walking score. Awkwardness in the performance of a motor act is often due to poor comprehension of the task given to perform; but after the intellectual counterpart of the motor act is learned, the resultant performance is quick and smooth. For example, a soldier with a high Rail-walking score (*e.g.*, 135 points) but a low mental age (*e.g.*, 9 years) will display in the first phase of training much difficulty and awkwardness in performing a proper right-about face. But after he has learned to place his right toe behind and slightly to the left of the left heel, to place his weight on his left heel and right toe and to swing himself about to the right, the act will ever after be done gracefully. Yet the soldier well within normal limits of intelligence but with a low Rail-walking score (*e.g.*, 45 points) will in all probability never perform an about face with smoothness and proper timing in spite of much extra training and attention. There is also evidence that the emotional factor merits consideration.

Another reservation in the general use of the test is the fact that it does not differentiate well at the upper levels of locomotor coordination. But qualitative observation of the test performance will aid in distinguishing those somewhat above average from those of superior coordination. The latter will hardly sway a bit, appearing to control balance by slight movements of the ankles, while the former in order to make a perfect score must depend upon grosser movements of hips and arms to maintain equilibrium. Consideration is being paid to the possible refinement of the measurement of this performance by use of electric eye beams parallel

to the rails and set at distances from the rail varying according to the broad or slender build of the subject. The make and break of the beam and resultant recording may lend itself to measurement of amplitude and frequency of body sway as well as supplying in graphic form a permanent record of test performance.

Finally the Rail-walking Test must not be considered when fine differentiation is needed. Even though the reliability of the test upon immediate reexamination is high ($r = .94 \pm .01$ for 100 subjects of wide range of score), from actual observation of Army training a significant difference in score between individuals is closer to 20 points than to 5 or 10 points. To date the test has proved useful in the Army as a gross instrument in diagnosing awkwardness and predicting trainability.

BIBLIOGRAPHY

1. HEATH, S. R., JR. Rail-walking performance as related to mental age and etiological type among the mentally retarded. *Amer. J. Psychol.*, 1942, 45, 240–247.

A COURSE IN ARMY PERSONNEL TECHNIQUES

BY CALVIN S. HALL
Western Reserve University

Shortly after Pearl Harbor, when it became apparent that the Army would have to expand its numbers greatly, the writer considered how best he might help to prepare undergraduate students with majors or minors in psychology to qualify for some form of psychological work in the Army. After some investigation of Army needs and available materials, a course in Army personnel meeting three hours a week for one semester was introduced at the beginning of the second semester of 1941–1942. It is being repeated each semester of the current year. Some of the students who completed the course last year and who are now in Army personnel work have reported that this course helped to secure them their present assignments.

In the belief that other psychologists may want to introduce a similar course, and because the materials are in some instances difficult to locate, the following outline of the course in Army Personnel Techniques is presented:

I. Introduction. A brief consideration of the meaning of personnel work, its objectives, obligations, and importance in contemporary society.

II. Historical Background. Personnel management in World War I. The principal source for this topic will be found in *The Personnel System of the United States Army*, published in 1919 by the Committee on Classification of Personnel in the Army (**27**), which unfortunately is now out of print. See also Yoakum and Yerkes, *Army Mental Tests*, 1920 (**30**).

III. German Military Psychology. The transition from the previous topic to this one is made by way of Dodge's prophetic statement written in 1919 (**14**):

> At the conclusion of our war work, two real dangers confront us, one military, the other, social. The military danger is that with the passing of the military crisis we shall stop our study of the mental factors in war. If some other country with more permanent policies should take up the

mental analyses where we have left them, and develop a real military psychology, they would have a military instrument vastly more effective than 42 cm. guns.

The discussion of German military psychology is based primarily upon Farago's monograph, *German Psychological Warfare* (**15**). See also the special number of the *Psychological Bulletin*, devoted to Military Psychology (**18**).

IV. Personnel Administration in the Army Today. The organization of the Army and the agencies charged with formulating and administering personnel administration are briefly discussed. This section of the material is taken largely from Van Riper's "Personnel Administration in the United States Army" (**29**). See also *The New Army of the United States*, a publication of the War Department (**26**). For the role played by psychologists in the Armed Services, see the *Journal of Consulting Psychology* for Sept.-Oct., 1941 (**20**), and the "Psychology and the War Section" of the *Psychological Bulletin* (**7**).

V. Interviewing: The Soldier's Qualification Card. The Soldier's Qualification Card (**23**) is the most important personnel record kept by the Army. It is a card about eight by eleven inches in size, both sides of which are used. Information concerning the soldier's family background, education, occupational experience, and other personal data are entered by an interviewer usually at a Reception Center. The scores of all tests administered in the Army are entered, and a tentative classification of the man is made by the Classification Specialist. The card is kept up to date by entering each change in the soldier's status.

During this section of the course the students are instructed in interviewing techniques and the proper completion of the Soldier's Qualification Card. Each student is required to buy a copy of *Army Regulations No.* 615–25, *Enlisted Men: Initial Classification*, dated July 31, 1942 (**2**); this pamphlet supersedes the one dated September 3, 1940. Facsimiles of the Soldier's Qualification Card are found in the back. Reference copies of the *Interviewer's Handbook* (**25**) published by the Personnel Procedures Section of the Adjutant General's Office are available, but much of the material found therein has been included in the preceding pamphlet, AR 615–25. For women students who are contemplating enlisting in the WAACs, there is a supplement to AR 615–25 entitled *Instructions for Women's Army Auxiliary Corps Initial Classification* (**3**).

Optional materials are *Army Regulations No. 605–90, Officer's and Warrant Officer's Classification* dated May 21, 1942 (1); mimeographed copies of *Officer's and Warrant Officer's Qualification Card;* and Bingham and Moore, *How to Interview*, 1941 (5).

VI. Trade Tests. In this section the student is familiarized with trade tests, their construction, administration, scoring, interpretation and uses in the Army. The *Manual of Oral Trade Tests* (16) used by the United States Employment Service is not available for distribution; but there are a number of trade tests reproduced in Chapman, *Trade Tests*, 1921 (12), and in Thompson, *Interview Aids and Trade Questions for Employment Offices*, 1937, (28). Students are also referred to Stead and Shartle, *Occupational Counseling Techniques*, 1940 (22), regarding the construction of trade tests, and they practice administering and scoring a number of such tests.

VII. Intelligence Tests. The Army General Classification Test is given to all soldiers. It is the most important single test used in the Army. Since the GCT cannot be discussed in class even if one should happen to be familiar with its contents, the Army Alpha is used as a substitute. The meaning of intelligence tests, their administration, scoring and interpretation, and their uses in the Army are discussed. Materials and references are: copies of Army Alpha; Yoakum and Yerkes, *Army Mental Tests*, 1920 (30); *Memoirs of the National Academy of Sciences, Psychological Examining in the United States Army*, vol. 15, 1921 (19).

VIII. Aptitude Tests. The two aptitude tests most frequently used by the Army are clerical and mechanical, although other special tests are used by the Army Air Forces, the Armored Force, and the Signal Corps. Since copies of these tests are strictly confidential, the Minnesota Clerical Ability Test and the Revised Minnesota Paper Formboard Test are used for class demonstration. The following topics are discussed in relation to aptitude tests: construction, administration, scoring, norms, reliability, validity, relation of scores to other variables, and critical evaluations of representative aptitude tests. Each member of the class selects or is assigned a test (not always an aptitude test) on which he must write a term report covering the topics listed in the first paragraph of this section. Materials and references are: copies of the two Minnesota tests, and files of various aptitude tests: Bingham, *Aptitudes and Aptitude Testing*, 1937 (6); Buros, *The Mental Measurements Yearbooks* (8, 9, 10, and 11); Greene, *Measurements of Human Behavior*, 1941 (17).

IX. Classification. The primary purpose of interviewing and testing is to provide suitable material on which to base a judgment as to the best assignment for each man in the Army. Proper classification is the crux of good personnel work and its importance cannot be over-emphasized. Materials and references are: the 720 page *Army Regulations, No. 615–26,* entitled *Enlisted Men: Index and Specifications for Civilian and Military Occupational Specialists, and Occupational Specifications for Non-English Speaking Men, Illiterates, and Men of Limited Mental Capacity,* of September 15, 1942 (**4**); the two-volume *Dictionary of Occupational Titles* prepared by the United States Employment Service (**13**); *Army Regulations No. 605–90* (**1**), regarding classification of officers.

X. Statistical Procedures. By using prepared statistical forms and adopting standard procedures for computing statistical constants, students learn to understand and compute the following: mean and standard deviation using grouped data, standard or z scores and percentile ranks, correlation using z scores, regression equations, standard error of estimate, and coefficient of alienation.

BIBLIOGRAPHY

1. Army regulations 605–90. Officer's and Warrant Officer's Classification. Washington, D. C.: U. S. Gov't Printing Office, 1942.
2. Army Regulations 615–25. Enlisted Men; Initial Classification. Washington, D. C.: U. S. Gov't Printing Office, 1942.
3. Army Regulations 615–25, Supplement. Instructions for Women's Army Auxiliary Corps Initial Classification. Washington, D. C.: U. S. Gov't Printing Office, 1942.
4. Army Regulations 615–26. Enlisted Men: Index and Specifications for Civilian and Military Occupational Specialists, and Occupational Specifications for Non-English Speaking Men, Illiterates and Men of Limited Mental Capacity. Washington, D. C.: U. S. Gov't Printing Office, 1942.
5. BINGHAM, W. V., & MOORE, B. V. How to Interview. New York: Harper, 1941.
6. BINGHAM, W. V. Aptitudes and Aptitude Testing. New York: Harper, 1937.
7. BRITT, S. H. (Ed.). Psychology and the war. *Psychol. Bull.*, April, 1942, to date.
8. BUROS, O. K. Educational, psychological and personality tests of 1933, 1934 and 1935. Stud. in Educ., No. 9, *Rutgers Univ. Bull.*, 1936, **13**, No. 1.
9. BUROS, O. K. Educational, psychological and personality tests of 1936. Stud. in Educ., No. 11, *Rutgers Univ. Bull.*, 1937, **14**, No. 2A.
10. BUROS, O. K. The 1938 Mental Measurements Yearbook. New Brunswick: Rutgers Univ. Press, 1938.
11. BUROS, O. K. The Nineteen Forty Mental Measurements Yearbook. Highland Park, N. J.: The Mental Measurements Yearbook, 1941.
12. CHAPMAN, J. C. Trade Tests. New York: Henry Holt and Co., 1921.

13. Dictionary of Occupational Titles. 2 vols., U. S. Employment Service, Washington, D. C.: U. S. Gov't Printing Office, 1939.
14. DODGE, R. Mental engineering during the war. *Amer. Rev. of Reviews*, 1919, **59**, 504–508.
15. FARAGO, L. German Psychological Warfare, Survey and Bibliography. New York: Committee for National Morale, 1941.
16. Manual of Oral Trade Tests. U. S. Employment Service, Washington, D. C.: U. S. Gov't Printing Office. (Not available for distribution.)
17. GREENE, E. B. Measurements of Human Behavior. New York: Odyssey Press, 1941.
18. PRATT, C. C. (Ed.). Military psychology. *Psychol. Bull.*, 1941, **38**, 309–508.
19. Psychological Examining in the United States Army. Memoirs, National Academy of Sciences, 1921, Vol. 15.
20. Psychological Service in National Defense. *J. consult. Psychol.*, 1941, **5**, 205–248.
21. Sound film on Classification of Enlisted Men. War Department, Serial No. 12–578. (May be obtained on loan from Central Distribution Library located at the Headquarters of each Service Command.)
22. STEAD, W. H., & SHARTLE, C. L. Occupational Counseling Techniques. New York: American Book Co., 1940.
23. The Adjutant General's Office. Soldier's Qualification Card, Form 20. Washington, D. C.: War Department.
24. The Adjutant General's Office. Officer's and Warrant Officer's Qualification Card, Form 66-1. Washington, D. C.: War Department.
25. The Adjutant General's Office. Interviewer's Handbook. Washington, D. C.: War Department, 1942.
26. The New Army of the United States. Washington, D. C.: War Department, 1941.
27. The Personnel System of the United States Army. 2 vols. Committee on Classification of Personnel in the Army, Washington, D. C.: U. S. Gov't Printing Office, 1919. (Out of print.)
28. THOMPSON, L. A., JR. Interview Aids and Trade Questions for Employment Offices. New York: Harper, 1937.
29. VAN RIPER, P., JR. Personnel administration in the United States Army. *Publ. Person. Rev.*, 1941, **2**, 199–210.
30. YOAKUM, C. S., & YERKES, R. M. Army Mental Tests. New York: Henry Holt, 1920.

THE COMMITTEE ON FOOD HABITS

BY MARGARET MEAD

Division of Anthropology and Psychology, National Research Council

The Committee on Food Habits was established in 1940 under the Division of Anthropology and Psychology of the National Research Council and operates under a contract with the Nutrition Division of the Office of Defense Health and Welfare Services. It parallels the Food and Nutrition Board, which operates under the same contract, under the Division of Biology and Agriculture. The function of the Committee is to formulate recommendations to the Director of Nutrition upon problems relating to food habits encountered in the coordination program of the Office of Defense Health and Welfare Services, and, where necessary to this end, to conduct studies and investigations and prepare reports upon such problems.

During 1941, the Committee had no substantial funds and confined its operations to occasional meetings and to the organization of two large conferences in May and June (**2**). In February, 1942, under the chairmanship of Dr. Carl E. Guthe, a permanent office was established in the National Research Council, with Dr. Margaret Mead as Executive Secretary and with two technical assistants, Dr. Natalie Joffe and Mrs. Rhoda Metraux.

The Committee on Food Habits is drawn mainly from anthropology, with single members representing the disciplines of medicine, nutrition, psychology, public health, and sociology. Members of the Committee are: Dr. Ruth Benedict, Dr. John M. Cooper, Dr. Allison Davis, Dr. J. K. Folsom, Mr. Lawrence K. Frank, Dr. Carl E. Guthe, Dr. Mark A. May, Dr. Nathan Sinai, Miss Mary E. Sweeny, Dr. Warren T. Vaughan, and Dr. W. Lloyd Warner. The Committee holds meetings at approximately two-month intervals, which include liaison sessions attended by the representatives of the various government bureaus and agencies whose interests in nutrition are coordinated by the Nutrition Division of the Office of Defense Health and Welfare Services. Problems of current importance are discussed, such as rationing, industrial feeding, and food and morale.

The Committee is officially charged with the task of providing expert opinions upon problems presented by the Nutrition Division and with the implicit obligation of developing and systematizing available scientific knowledge on the subject of food habits. During the past year the Committee has undertaken a variety of tasks:

1. It is developing a systematic approach to the problem of food habits, with particular reference to food habits of the American people at the present time.

2. On selected sub-cultural groups it prepares memoranda organized to express the point of view of the Committee, and it gives these memoranda limited professional circulation (**1**, **4**, **5**, **6**, **7**, **18**).

3. Tests have been developed which can be used with school children in different parts of the country to reveal the characteristic food ideologies of different age levels and of different economic, occupational, religious and national background groups. This work has been done by Professor Kurt Lewin (**9**) and Mr. Alex Bavelas at the Child Development Station, State University of Iowa, and the tests have been validated on a sample of some 2,300 school children in Cedar Rapids, Iowa. Extensive background materials, including intensive interviewing and use of questionnaires, accompanied the construction of these tests.

4. Controlled experiments (**10**) have been made in methods of influencing food habits in real life situations at the State University of Iowa and under the direction of Mr. Earl L. Koos (**8**) with the sponsorship of Cornell University Medical College, Department of Public Health. In these experiments, directed toward determining the relative usefulness of the block plan as compared with the utilization of natural "friendship" groups, it was demonstrated that friendship groups were unsuitable for the dissemination of nutrition information—at least, among low income urban people of foreign background.

5. A series of current reports has been compiled for administrative use, dealing with American attitudes toward the nutrition program, meat shortages, rationing, hoarding, etc. These reports are based upon verbatim materials collected by undergraduate and graduate students in attitude sampling squads in cooperating universities and colleges, and analyzed by Mrs. Rhoda Metraux. The results have been supplemented in some cases by the researches of individual women psychologists, cooperating with the Committee through the Subcommittee on Services of Women Psychologists in the Emergency. In the analysis of the verbatim materials there have been two methodological innovations: (a) a method by which groups of students in the universities throughout the country could make a definite contribution to the war effort, and (b) a new method of analysis summarized in a graphic presentation, in a form suitable for administrative use, of the structure of the attitudes revealed in the verbatim materials. This method is based upon the anthropological assumptions that culture is systematic, and that by relating the regularities in responses of different groups all over the country, reliable clues to American attitudes may be found which are of a different order from those obtained by opinion sampling. A summary of these studies (**16**) is available for limited distribution.

6. The Committee has sponsored long-time research, aided by grants from private institutions, which contribute to our understanding of food habits. Two such projects are now under way: (a) a study concerning the psychosomatic mechanisms in obesity and allergy in childhood and attitudes involved in children's changing food habits, conducted by Dr. Hilda Bruch at The Johns Hopkins University, and (b) a study of the influence on food habits of different containers, by Professor Kurt Lewin of the State University of Iowa.

7. The Committee has sponsored and assisted in the compilation of pieces of voluntary research undertaken by graduate students on subjects immediately relevant to the wartime nutrition program (**3**, **17**).

8. It has served as a clearing house for food-habits projects, both in government agencies and throughout the country.

9. It has provided consultation and collaboration for those conducting research on food habits.

10. Through its Executive Secretary the Committee has been called upon for materials for current publications bearing on food habits (**11**, **12**, **13**), and for addresses at scientific meetings on the relationship between food-habit research and current problems (**14**, **15**).

In research the Committee is interested in experiments and studies of: (1) the content of food habits of special groups in the population, by locality, age, sex, national background, or race; (2) the motivations involved in food habits and any attempts to use these motivations for change; (3) different methods of changing food habits and methods of evaluating these experiments; (4) conditions involving food habits, e. g., asthma, diabetes, obesity, anorexia nervosa; (5) feeding problems and integration of eating habits with personality attitudes; (6) nutritional status as correlated with social or psychological factors such as education, socio-economic status, occupation, etc.; (7) the relationship between the emerging food problems connected with hoarding, fear of rationing, and rising prices and general morale; (8) techniques which have proved efficacious in introducing nutritional knowledge into community or agency programs; (9) the relationship of idiosyncratic physiological status to food choice; (10) experiences with reference to group feeding, school lunches, emergencies such as fire or flood, soup kitchens during the depression, etc.; (11) the relationship between food habits and other habits such as punctuality, use of time, etc.; (12) the way in which giving or withholding of food is used as a sanction in bringing up children; (13) the role of the nutritionist or home economist in different types of community program; and (14) changes in food habits as they provide leverage for other types of social change.

BIBLIOGRAPHY

1. Committee on Food Habits, National Research Council. The relationship between food habits and problems of wartime emergency feeding. May, 1942 (mimeographed).
2. Committee on Food Habits, National Research Council. Summaries of the 1941 conferences on research in the field of food habits, contributions from the field of market research, contributions from the field of child development. May–June, 1941 (mimeographed).
3. FRISCH, G. E. A study of the effects of odd-shifts upon the food habits of war workers. Committee on Food Habits, National Research Council, August, 1942 (abstract mimeographed).
4. JOFFE, N. F., & NIZZARDINI, G. Italian food patterns and their relationship to wartime problems of food and nutrition. Committee on Food Habits, National Research Council, August, 1942 (mimeographed).
5. JOFFE, N. F. Hungarian food patterns and their relationship to wartime problems of food and nutrition. Committee on Food Habits, National Research Council, February, 1943 (mimeographed).
6. JOFFE, N. F., & BENET, S. Polish food patterns and their relationship to wartime problems of food and nutrition. Committee on Food Habits, National Research Council, February, 1943 (mimeographed).
7. JOFFE, N. F., & PIRKOVA-JAKOBSON, S. Czech food patterns and their relationship to wartime problems of food and nutrition. Committee on Food Habits, National Research Council, February, 1943 (mimeographed).
8. KOOS, E. L. Food in the lives of our neighbors. The District Health Committee, Kips Bay-Yorkville Health District, New York City, August, 1942, 1–47.
9. LEWIN, K. A group test for determining the anchorage points of food habits. Committee on Food Habits, National Research Council, July, 1942 (mimeographed).
10. LEWIN, K. The relative effectiveness of a lecture method and a method of group decision for changing food habits. Committee on Food Habits, National Research Council, July, 1942 (mimeographed).
11. MEAD, M. The factor of food habits. *Ann. Amer. Acad. Polit. & Soc. Sci.*, 1943, **225**, 136–141.
12. MEAD, M. Reaching the last woman down the road. *J. Home Econ.*, 1943, **34**, 710–713.
13. MEAD, M. Changing food habits. Report of the New York State Joint Legislative Committee on Nutrition (in press).
14. MEAD, M. Dietary patterns and food habits. *J. Amer. Diet. Assoc.* (in press).
15. MEAD, M. Psychoanalytic contributions to the problem of changing food habits. *Bull. of Menninger Clinic* (in press).
16. METRAUX, R. Qualitative study of current attitudes on food problems. Committee on Food Habits, National Research Council, August, 1942 (mimeographed).
17. METRAUX, R. Rationing and morale. Committee on Food Habits, National Research Council, September, 1942 (mimeographed).
18. SWEENY, M. E. Changing food habits. *J. Home Econ.*, 1942, **34**, 457–462.

THE ATTITUDES OF YOUTHS OF HIGH SCHOOL AGE TOWARD THE WAR

BY MANDEL SHERMAN
University of Chicago

During June and July, 1942, a study was made of the attitudes toward our war effort of 7,000 Chicago youths of high school age. The attitudes were obtained from English themes of 6,500 youths and from personal interviews with 500 other youths, the majority of whom were not attending high school.

The themes were written during the regular English periods at the direction of the regular teachers. The subject of the themes was "How the War Affects Me." No other instructions were given to these students, and they were asked to write as rapidly and as fully as possible. It was felt that the attitudes of these students would be expressed more freely if they wrote as a regular class assignment, as compared with writing or talking about them in some other situation.

It is obvious that it is difficult to obtain the attitudes of youths or adults because of various defensive reactions and because rapport between the interviewer and the interviewee is difficult to establish. A common method of discovering attitudes towards public problems is to ask direct questions. Usually, however, persons who are interviewed are quite guarded in their responses, and are likely to give answers which they believe are expected of them. Furthermore, it is exceedingly difficult to anticipate the problems involved in shaping the attitudes of an individual, and thus the ordinary questionnaire is not effective in obtaining the attitudes of the majority of those interviewed.

The title "How the War Affects Me" was chosen in order to avoid specific questioning on any one item or element relating to the war. As was expected, the subjects wrote on a variety of topics, often unrelated to the effects of the war upon themselves. The most serious drawback to this method of investigating the attitudes towards the war was the variation in the length of the themes which these students wrote for their teachers. Some students wrote three

to four pages, whereas others wrote a page, and a small number wrote as little as a half page. It is also possible that the introduction of the topic was too sudden and that preparation for the writing of the theme might have influenced many of the students to write longer papers. These undesirable elements were compensated for, however, by the neutral attitude of the teacher and the fact that the students expected that they would be judged mainly on their ability in writing the theme.[1]

The papers of these students were read by three persons in addition to the investigator. Before ratings were made, discussions were held regarding the purposes of the study, and samples of the papers were read aloud in order that there should be mutual agreement upon the various ideas expressed in the papers. Each judge then placed the papers he read in one of seven groups. A sampling was taken of each of these groups and the following classifications of the general attitudes of these subjects were made:

1. Direct antagonism to the war, including criticism of the government, of the war effort, and of the reasons for the war (more emotional and general than classification 2). *5 per cent.*

2. A critical attitude towards the war and towards the government, including criticism of the war and how the war is handled, criticism of rationing, etc. *6 per cent.*

3. Indifference. In this category statements were made that the war "perhaps could not be avoided," or "since we are in the war, we might as well go through with it." *21 per cent.*

4. Confusion, where the youth states that the whole war is a sort of a mystery and he does not know what to think. Often statements are made that perhaps it might be all right to have a war because people will learn how to save. *12 per cent.*

5. Mildly favorable attitude towards the war, but critical of specific elements. For example, one boy stated that there is no reason for taking young people, that older persons should go. Another boy said that they are going to draft his father and there is no reason why his family should be penalized because the nation goes to war. *9 per cent.*

6. A favorable attitude towards the war with the statement that he will try to help as much as possible, for example, by selling stamps, or by becoming a nurse or an airplane pilot. *26 per cent.*

[1] In discussing this research project with several colleagues, they pointed out that when the students were expected to write on "How the War Affects Me" they might have thought in adverse terms. In other words, they may have interpreted the question to mean that they must write on the adverse effects of the war, rather than on their true attitudes. As will be shown in this paper, this is a possibility but not a probability because of the different reactions of these students and the fact that such a large number of subjects did not make any mention of the adverse effects of the war upon them.

7. A strong favorable attitude with, perhaps, statements that we should have gotten in earlier and with parallel vindictiveness against the enemy. *21 per cent.*

Those who were against the war effort (classification 1) and who were 5% of the total group, frequently not only expressed an attitude that the Germans or Italians were justified in their aims, but also expressed an antagonism towards the President's policies.

Nineteen per cent of these youths stated that the Germans were forced to fight, and of these, over 90 per cent also expressed the attitude that the policies of this government resulted in our getting into the war. Thirty-five per cent of those who stated the Germans were justified in beginning the war also stated that we were responsible for our participation in the war by goading the Japanese to attack Pearl Harbor, or by conniving to get us into the war in order to profit special persons. The following excerpts from the papers of these youths are examples of their antagonistic attitudes:

"I think this was brought upon us because we did not know how to conduct ourselves with other nations. People say our enemies were aggressors, but we were aggressors also."

"If I were the President, I would never have sent so much material to England so that we now have to fight for England. I am not excusing Germany or Japan, but we could have made friends with everybody."

"War is a terrible thing and we should have done everything to stay out of it. Maybe Germany was taking advantage of us, but other countries have also taken advantage of us."

"I am as patriotic as anyone, but I believe that I am honest when I say that the President and the men around him goaded Japan to attack us. This gave him an excuse to get into the World War."

Those youths who were critical towards the war effort but did not show direct antagonism (classification 2), who were 6% of the total group, most frequently expressed the belief that this war is wrong because it will solve nothing and that we would not have gotten into the war had we conducted ourselves differently. They believe that there will be a serious depression after the war. Therefore, the war not only would not solve anything but would also leave us worse off than ever before. Thus these youths not only criticized the government, but expressed resentment against present and future deprivations.

Approximately 10 per cent of these subjects stated that the treaty after the last war was the direct cause of this war. The following are examples of the statements made and the items mentioned by students whose attitudes were placed in classification 2:

"I think the war was uncalled for in the first place, but I suppose we will have to go through with it since we are in it. When I read in the paper that they have so much sugar in the warehouses that they don't know

what to do with it and to think of all the millions of dollars the Government spent for it, it makes things look fishy."

"Maybe the war is all right because we had to get into it sometime, but we should not get in too good with Russia. They are also the kind of people you cannot trust."

"I really don't think we should have gotten into the war in the first place. Maybe we will come out of it all right and maybe not, but I am not sure."

"This war is not going to do us much good even though we are supposed to be fighting for democracy. As far as shipping materials and guns and food across the sea, that is all right, but I think everything we ship them should be paid for right away."

The statements of the youths which showed an indifferent attitude were placed in classification 3 and constituted 21% of the total group. A large number of these stated that the war perhaps could not be avoided, or since we are in the war we might as well go on with it. They seemed to express a somewhat neutral attitude, with statements that the war will be a long one and there might be a depression after the war.

About 40 per cent of the attitudes in classification 3 included statements about deprivation, such as a depression after the war, their inability to go to college, statements that there was no use saving money because inflation would surely occur, and complaints that rationing is to be instituted because of shortages and also because the government wanted to make us war conscious. The following are brief excerpts to illustrate these attitudes:

"This war doesn't mean much to me except that we have to get along with less sugar and maybe other things. But we have everything at home."

"I suppose we might as well get it over with because things might be different later."

When a subject showed confusion regarding the war and its issues, his attitudes were placed in classification 4. These constituted 12% of the total.

About a fifth of these youths stated that the war is a mystery to them. Over 30 per cent stated that in spite of all that they read about the war they really did not know what we are fighting for. Over 50 per cent emphasized their loyalty, but also emphasized their confusion about the war effort. Approximately 20 per cent stated that they had talked with their parents or other boys and girls about what the war meant to them, and they had not made up their minds as yet.

The 9% of the youths who showed a mildly favorable attitude towards the war effort neutralized, in some degree, by a variety of criticism were placed in classification 5. The attitudes were put

into this classification because they were definitely unsatisfied with one or more elements.

For example, one eighteen-year-old stated that there was no reason for drafting young people, and another, that rationing was unnecessary. Thirty per cent of this group stated that the administration in Washington was becoming dictatorial and was concerned more with its power than with winning the war. Approximately 10 per cent complained because there was too much censorship. Over 15 per cent stated that they were fairly sure that most of the navy was sunk but that the newspapers were not allowed to print the news. These youths in general said that they approved of the war but did not elaborate their statements. They merely stated that the dictators must be curbed, that war is inevitable in a world at war, that we were not quite smart enough to stay out of the war, but that we cannot expect to have supermen. Approximately 25 per cent of this group stated that they approved of the war principally because this conflict is so serious that people will learn to stay out of wars in the future. Not one of this group mentioned the seriousness of the possibility of losing the war, or the threat to us by the aggressors, nor did they mention any positive issues, such as the four freedoms or democracy.

Those who showed a favorable attitude towards the war without specific criticism of various war efforts were classified in group 6. Twenty-six per cent of the total number of subjects were thus classified.

These individuals did not mention rationing except that approximately 9 per cent complained about the attitudes of some people who are dissatisfied with rationing. Fully 50 per cent stated that we should have gotten into the war earlier and over 40 per cent stated that they felt sure that we will win the war.

The attitudes which were strongly favorable towards the war effort and which also included a vindictive attitude against the enemy were classified in group 7 and accounted for 21 per cent of the total.

Attitudes regarding their personal deprivations were included in approximately 53 per cent of the responses of the youths who were not strongly in favor of the war. The most consistent feeling of deprivation revolved about rationing and about the depression. Thirty-two per cent of these youths believed that a depression would follow the war, and about 75 per cent believed that more and more rationing would occur quickly. Eight per cent believed that we are headed for Socialism and Communism. Only 25 per cent made specific mention of the outcome of the war.

Of the youths whose attitudes were classified in the first four categories (44 per cent of the total), *i.e.*, antagonism, critical atti-

tude towards the war, indifference, and confusion, 65 per cent expressed various fears regarding the war. Of these, 8 per cent expressed the fear that the economic depression will be more serious than it ever has been previously; 10 per cent expressed the fear that everything would be ruined, that inflation would wipe out all savings and that no one will be able to get a job; 15 per cent stated that the suffering during and after the war will be intense; and 32 per cent expressed a general fear about the killing going on in the war, and their fears of seeing wounded soldiers.

When the contents of the students' papers were analyzed, more verbalizations were found in the higher age groups. Students in the third and fourth year of high school also showed approximately 10 per cent greater frequency of criticism either of the war or of governmental agencies than did younger students. The girls at every age group were more verbal than the boys and showed stronger attitudes either for or against the war effort, and in relation to the criticism of given war activities.

Summary

1. Of 7,000 Chicago youths of high school age, who were studied with respect to their attitudes toward the war, it was found that: 5 per cent showed direct antagonism toward the war, 6 per cent had a critical attitude, 21 per cent were indifferent, 12 per cent were confused, 9 per cent were mildly favorable, 26 per cent were favorable, and 22 per cent were strongly favorable.

2. In youths not strongly favorable toward the war, strong attitudes with regard to deprivations were found in 53 per cent of the cases.

3. In youths not favorable to the war in any degree, 65 per cent expressed various fears regarding the war.

4. There was more verbalization and more criticism in the older youths.

5. Girls were more verbal than boys and showed stronger attitudes.

PSYCHOLOGY AND THE WAR: NOTES

Women Psychologists and the Armed Services. Inquiries have been received from women psychologists as to the possibilities for psychological service within the Women's Army Auxiliary Corps (WAAC); Women Appointed for Voluntary Emergency Service (WAVES); the Women's Reserve of the Coast Guard (SPARS—semper paratus—always ready—motto of the Coast Guard), and the Marine Corps. Although it is not possible to give a definitive answer to this question, plans at headquarters of these organizations suggest that the services of women psychologists can and will be used both among officers and enlisted personnel. Psychological training will be particularly useful with reference to problems of allocation and placement of personnel.

A wide range of activities is to be carried on by the members of the three major groups. A composite list of activities includes:

*Accountants
*Aircraft Specialists
  Electrical
  Instrument
  Power Plant
  Propeller
*Bakers
#Bandsmen
#Bookbinders
*Bookkeepers
*Cashiers
#Classification Specialists
*Clerks
  General
  Commissary
  File
  Library
  Medical
  Postal
  Record
*Control Tower Operators
*Cooks
##Dental Assistants
*Dietitians
*Draughtsmen
*Drivers (automobile)
*Electricians
#Glider Instructors
##Hospital Assistants
#Instructors
##Laboratory Assistants
*Librarians
*Link Trainer Instructors
*Machinists
*Messengers
*Metal Workers
*Mimeograph Operators
*Motion Picture Projectionists
##Pharmacists
*Photographers
*Photo Laboratory Technicians
##Printers
*Radio Operators
*Radio Technicians
*Receptionists
*Secretaries
*Statisticians
*Stewardesses (mess sergeants)
*Store Managers
*Telegraph Operators
*Telephone Operators
*Teleprinters
*Teletype Operators
*Teletype Repairmen
*Typists
##Waitresses
*Weather Observers
#Woodworkers

All women who enroll in any of the services are urged to notify the Office of Psychological Personnel, 2101 Constitution Ave., Washington, D. C., immediately, giving full details. Insofar as possible steps will be

* Common to all services
\# WAAC'S only.
\#\# WAVES and SPARS only.

taken to facilitate their functioning as psychologists. Following is a brief summary of enlistment procedures and conditions of service and training.

WAAC. The Women's Auxiliary Army Corps is organized for non-combatant service with the Army, to take the places of men in non-combatant positions and release them for front-line duty. Enrollment must be as an auxiliary. After four weeks of basic training, aptitude tests are administered to determine special fitness, after which further training at a specialist school may be indicated. The recent expansion of the Corps affords a wide range of opportunities for advancement of enrolled members, both as non-commissioned officers and through promotion to commissioned rank. All officer candidates are selected from the ranks of enrolled members to attend Officer Candidate School. At the present time the officers in the Corps with previous experience in personnel work are being used in selection and placement duties. The length of service is for the duration, plus a period of not more than six months thereafter. An applicant must be a citizen between 21 and 45, who can submit satisfactory proof of date of birth and citizenship and two character references. She must be able to qualify physically and to pass a mental alertness test. Married women can serve if they have no financial dependents. The monthly pay of an enrolled member is:

Auxiliary	$ 50.00
Auxiliary, 1st class	54.00
Junior Leader	66.00
Leader	78.00
Staff Leader	96.00
Teachnical Leader	114.00
First Leader	138.00

Officers' monthly pay is:

Third Officer	$125.00
Second Officer	131.25
First Officer	166.67

Further details may be obtained from any United States Army Recruiting and Induction Station.

WAVES and SPARS. The WAVES is an organization whose members replace Navy men at shore stations; the SPARS replace Coast Guard men at shore stations. Duties include all types of office work, jobs in radio, communications, storekeeping, and various special and technical positions. Basic training of enlisted personnel averages four months in a training school at one of the selected colleges over the country. The term of enlistment is the duration, plus no more than six months.

Enlisted WAVES or SPARS must be physically qualified citizens between 20 and 35, with at least two years of high school or business school, able to furnish three character references. A married woman may enlist in the WAVES provided her husband is not in the Navy, in the SPARS if her husband is not in the Coast Guard. Officers must meet the same requirements, except that they may be from 20 to 49 years of age, must have a college degree or two years of college plus at least two years of

acceptable business or professional experience. Also, two years of mathematics in high school or college is required for the WAVES. All officer candidates receive preliminary training at Smith College. (Incidentally, the head of the SPARS is a woman psychologist, Lt. Commander Dorothy Stratton.)

Monthly base pay of enlisted WAVES or SPARS is:

Apprentice Seaman	$ 50.00
Seaman, Second Class	54.00
Seaman, First Class	66.00
Petty Officer, Third Class	78.00
Petty Officer, Second Class	96.00
Petty Officer, First Class	114.00
Chief Petty Officer, Acting Appointment	126.00

Officers draw the same pay and allowance as men officers of equivalent rank in the Navy:

Ensign	150.00 plus allowances
Lieutenant, Junior Grade	166.67 plus allowances
Lieutenant, Senior Grade	200.00 plus allowances

Additional information may be secured from any branch of the Office of Naval Officer Procurement:

Atlanta, Ga. . . . . Healy Building
Baltimore, Md. . . . . Richmond Market Armory, North Howard Street
Birmingham, Ala. . . . . 601–609 Jackson Building
Boston, Mass. . . . . 150 Causeway Street
Buffalo, N. Y. . . . . Liberty Bank Building
Charleston, S. C. . . . . The Center, Marion Sq.
Chicago, Ill. . . . . Board of Trade Bldg., 141 West Jackson
Cleveland, O. . . . . 6th Floor, Marshall Building
Columbia, S. C. . . . . University of South Carolina
Dallas, Tex. . . . . 1530 Allen Building
Des Moines, Ia. . . . . 2nd Floor, Old Post Office Building
Detroit, Mich. . . . . 9th Fl., Book Bldg., 1249 Washington Blvd.
Houston, Tex. . . . . 824 Niels Esperson Building
Indianapolis, Ind. . . . . 429 North Pennsylvania Street
Jacksonville, Fla. . . . . 915 Lynch Bldg.
Kansas City, Mo. . . . . 202 Finance Bldg., 1009 Baltimore Ave.
Los Angeles, Calif. . . . . 850 Lilac Terrace, U. S. Naval & Marine Corps Res. Armory
Miami, Fla. . . . . Room 905, Langford Bldg., 121 S. E. 1st St.
Minneapolis, Minn. . . . . 4th Fl., Roanoke Bldg., 109 South 7th St.
Nashville, Tenn. . . . . Third National Bank Bldg., 3rd and Church
New Orleans, La. . . . . Louisiana Bldg., 217–227 Camp Street
New York, N. Y. . . . . 33 Pine Street
Norfolk, Va. . . . . 425 Federal Building
Oklahoma City, Okla. . . . . Post Office Building
Philadelphia, Pa. . . . . 17th Fl., Widener Bldg., Juniper & Chestnut Sts.
Pittsburgh, Pa. . . . . Keystone Hotel Bldg., 3rd Ave., & Wood Street

Portland, Ore. ..... 1233 American Bank Building
Raleigh, N. C. ..... North Carolina State College
Richmond, Va. ..... 2nd Fl., Chevrolet Parts Bldg., Norfolk & Altamont
San Francisco, Calif. ..... 703 Market Street
Seattle, Wash. ..... 117 Marion Street
St. Louis, Mo. ..... 6th Fl., Missouri Pacific Bldg., 210 N. 13th St.
Washington, D. C. ..... 1320 G Street, N.W.

Marine Corps. The Marine Corps has established a women's reserve to be known simply as marines. Women between 20 and 50 are eligible as officer candidates, and between 20 and 36 as enlisted personnel. Applicants for commissions must have a college degree, or two years of college and at least two years of business experience, or "special qualifications in a particular field." Enlisted women must have at least two years of high school or business school, or special qualifications. Pay and rank are identical to that of men officers and enlisted personnel. Details can be obtained from any of the Offices of Naval Officer Procurement listed above.

Psychologists in War Work. During December, 1942, Dr. Walter S. Hunter, Chairman of the Department of Psychology at Brown University, made a study of supply and demand with reference to psychological personnel, acting as a committee of one for the Emergency Committee in Psychology of the National Research Council. Certain aspects of the study are of general interest.

The National Roster of Scientific and Specialized Personnel, as of November 1, 1942, lists 2,927 and 991 women with master's or doctor's degrees in psychology, a total of 3,918 professional psychologists. The American Psychological Association and its affiliated societies, which contain practically all who have any claim to a professional status, include approximately 3,850 individuals of whom about 1,000 are women. The total number of professional psychologists in this country, at a liberal estimate, does not exceed 3,000 men and 1,000 women. A more conservative estimate is 2,500 men and 900 women. Of this total professional group, many are unavailable for service outside of their present positions because of present essential civilian war work, age, physical disability, economic factors, dependency, and the like. Of those actually available, many would need refresher courses and training on specific jobs before they could serve effectively.

At the present time at least 25 per cent of the total pool of professional psychologists are serving as psychologists in the Army, the Navy, and the civilian agencies of government.

Registration with Office of Psychological Personnel. Both men and women psychologists interested in war work or other placement opportunities should continue to register themselves with the Office of Psychological Personnel. Names presented to various agencies are primarily those of psychologists who have indicated their availability. Psychologists should also, of course, see to it that their names are on appropriate registers of the United States Civil Service Commission.

If employment, address, or Selective Service status changes at any time, *both* the Office of Psychological Personnel, and the National Roster of Scientific and Specialized Personnel should be notified immediately. The Office of Psychological Personnel is located in the National Research Council, 2101 Constitution Ave., Washington, D. C.; the National Roster is at 10th and "U" St. N.W., Washington, D. C.

Committee on War Services to Children. The Committee on War Services to Children of the American Psychological Association is a continuation of the Committee on Refugee Children which was appointed in 1940 with the following members: Dr. John E. Anderson, Dr. Gertrude Hildreth, Dr. Kathryn E. Maxfield, Dr. Catherine C. Miles, Dr. Robert T. Rock, Dr. Gladys C. Schwesinger, Dr. Mary M. Shirley, Dr. Frank K. Shuttleworth, Dr. Anna S. Starr, and Dr. Arthur T. Jersild, Chairman.

The Committee has concerned itself with two lines of activity: direct psychological service in connection with war efforts in behalf of children, and research or fact-finding. The Committee has offered its services to various national and regional agencies. Members of the Association who are in contact with war work as it relates to children are invited to inform the Committee concerning their activities and to suggest ways in which the Committee might be most useful.

On the research side, the Committee, through one of its members, is preparing, for publication in the *Psychological Bulletin*, a critical review of literature dealing with the impact of war on children. It is planned that further comprehensive surveys be prepared as occasion demands during and following the war. The Committee would appreciate word from psychologists who have studies under way or who know of studies that are in progress.

BOOK REVIEWS

GRIFFITH, JOHN Q. & FARRIS, EDMOND J. The rat in laboratory investigation. Philadelphia: Lippincott, 1942. Pp. xiv + 488.

Almost twenty years ago Dr. H. H. Donaldson, a member of the Wistar Institute, published his book, *The Rat*. That book was an almost indispensable reference book for anyone dealing with this animal in the laboratory. The present book, which bears many earmarks of the same Institute, is just as indispensable. The indispensability of the two books rests, however, upon different foundations. Dr. Donaldson's book is largely a book of norms of growth and development of the rat while the present manual is concerned chiefly with techniques of handling and studying the animal.

The *raison d'etre* of this manual is largely the tremendous use of the albino rat as a laboratory animal. No other species is so widely used. The rat is found in use in the laboratories of the physiologist, psychologist, biochemist, dietician, geneticist, and many more. It is probably safe to say that man knows more about the rat than he does about himself. It is one of the purposes of the thirty contributors to this book to describe the techniques by which this tremendous mass of information has been discovered. The value of the work is still further enhanced by the fact that many of these techniques may be applied to other animals as well.

The first two chapters are concerned with the general problems of the handling, breeding, and maintenance of health in the colony. The methods and apparatus described are chiefly those used in the large colony of the Wistar Institute. The material in these chapters will probably impress most readers as a description of ideals to be sought rather than a picture of the usual rat colony. Few laboratories have expended the money and care in the treatment of their rats that the Wistar Institute has. Nevertheless, the technician in even the most modest colony will find a number of very useful and practical suggestions.

Chapter 3 is a highly condensed version of the large and elaborate book by the same author on the anatomy of the rat. In 23 pages all parts of the rat's body are described in sufficient detail to serve as a foundation for much of the material that follows, particularly that concerned with surgical techniques. A number of drawings, several in colors, are used to indicate the main portions of the rat's skeleton and the mutual relationships of these parts.

Very little work has as yet been done with the rat embryo, but that it is feasible to do so within limits is indicated by a description of tried methods in Chapter 4.

Probably more experiments with the rat as a subject have been done on dietary deficiencies and requirements than on any other one subject.

The rat, being omnivorous, reacts to most food substances in much the same way as the human being. This accounts in a large measure for the animal's widespread use in diet laboratories. A large amount of this work has been summarized in Chapter 5. In the neighborhood of 60 different substances are discussed with reference to their role in the diet of the rat. The author refers to 269 journal articles.

The reader might normally expect to find chapter 6, on the teeth of the rat, in the section on the anatomical material. However, because of the peculiar properties of the tooth, "it is a valuable biologic indicator which mirrors and records, during its development, the metabolic status of the animal. This record, unlike similar records in the bones, is not transitory but is permanently engraved upon the enamel and dentine in which resorption does not occur." Therefore, an extensive discussion of the teeth is given with many diagrams and illustrations. Of particular interest is the description of the effect upon the teeth of various experimental factors such as diet, endocrine disturbances, etc.

Chapter 7, on the digestive system, describes a number of surgical techniques and the purposes for which they may be used. The author feels that experimentation on the digestive system of the rat has not been as extensive as it might well be and that there are many unexplored potentialities.

The metabolism of the rat is discussed in Chapter 8. Methods of measuring and factors affecting it are discussed.

Various surgical techniques for the investigation of the central nervous system occupy the few pages in Chapter 9.

By far the longest chapter in the book is the tenth, on "Technics for the Investigation of Psychological Phenomena in the Rat." This is perhaps a reflection of the fact that the white rat has been traditionally associated with the psychological animal laboratory. The technics range all the way from those dealing with tropistic behavior to social reactions. As might be expected in this chapter, there is noted a great amount of disagreement about what a given technic does or does not do. Much of this revolves around the lack of a standard terminology. To a neurologist, the term cisternal puncture indicates a definite and well-defined operation, but in psychology many of the terms used have almost as many significants as there are psychologists. This condition is caused by a number of factors among which are: (1) the newness of the discipline; (2) the complexity of the subject matter; (3) the taking over of many terms from common usage; (4) the inclination of some animal psychologists to use laymen's terms because of their attention-attracting value when applied to the animal.

On the other hand, it is probable that the standardization of experimental techniques can never go as far in animal psychology as it has in many other fields. This is because the psychology of even the rat is incomparably more complex than its digestive system, circulatory system, or any other part of its anatomy. Its psychology is the integration of all of these systems.

Chapter 11, "The Circulatory System," describes methods of taking blood pressure, taking blood, inducing pathologic states, etc.

The rather limited use which has been made of the rat in the biologic assay of hormones is outlined in Chapter 12.

On numerous occasions the reviewer has wished to have at his finger tips the information of Chapter 13. The greater part of the chapter is given over to a table showing the dosages and results of administration of a large number of drugs and chemicals. There are many factors which may help to determine the effects of a drug. Some of these are discussed and in the table the reference for each entry is given so the reader may refer to the original article for more details.

Chapter 14 considers the technics of blood study and blood diseases. A very interesting discussion is that of blood groups in the rat.

Chapter 15 is concerned with X-ray technic. The authors feel that the rat should be used more frequently in studies which require the making of roentgenograms because the size of the animals allows the use of portable and economical apparatus and small films. A number of interesting roentgenograms are reproduced.

Detailed directions for doing a number of surgical operations on the rat are given in Chapter 16. Perhaps the most difficult of the operations described is that of hypophysectomy. About 23 technics in all are outlined.

Histologic technics as applied to rat tissue are not particularly different from those used for other mammals. However, a number of formulae and methods which have been successfully used with rat tissue are given in Chapter 17.

The preparation of the skeleton for study is described in Chapter 18. Methods of chemical analysis are also indicated.

A very short discussion of the eye of the rat makes up Chapter 19. Most eyes have a hyperopia of about 5 diopters. Cerebellar tumors may affect the venous picture of the fundus.

Chapters 20 and 21 are concerned with animal parasites of the rat, the former with protozoa and the latter with metazoa. Some of these infestations are of great medical importance because of their part in spreading certain human diseases. The majority are important, however, only from a general scientific point of view and in so far as they affect the health and comfort of the rat under laboratory conditions.

The final chapter of the book will be of great interest to all who use rats for laboratory purposes. Nothing is quite so discouraging as having an infection wipe out a group of animals upon which a great deal of time and effort has been expended. This has happened once in the reviewer's experience, when the disease was paratyphoid. Although so high a degree of virulence of this disease is rather rare, it is not at all uncommon to find individual cases of it in a colony. Another disease which is very common and which often seriously incapacitates the animal even when it does not result in mortality is pneumonia. This is the disease which is often called "sniffles."

Cleanliness, adequate diet, plenty of fresh water, fairly constant temperature, and a sturdy stock will generally suffice to keep a colony healthy.

The writer has tried to give the reader of this review a very brief

indication of the contents of this book. The names of the authors have been omitted to save space. Some chapters are written by several authors and the same author may appear at several different places in the book. One has the impression as he studies the material that each author is writing on a subject with which he is thoroughly familiar. Documentation is very adequate.

The mechanics of the book are excellent. It is well printed, and glossy paper is used because of the presence of so many half-tones. It is bound in sturdy binding which looks as if it would stand being wiped off occasionally with a damp cloth. Thus, one may feel free to have the book at one's elbow in the laboratory.

The present reviewer can find no significant adverse criticism. His opinion is that the *The Rat in Laboratory Investigation* is a necessity for any experimenter or laboratory using the rat.

W. T. HERON.

University of Minnesota.

BOOKS AND MATERIALS RECEIVED

GELLHORN, E. Autonomic regulations. New York: Interscience Publishers, 1943. Pp. xii+373.

PENNINGTON, L. A., HOUGH, R. B., & CASE, H. W. The psychology of military leadership. New York: Prentice-Hall, 1943. Pp. ix+288.

RANSON, S. W. The anatomy of the nervous system. (7th Ed. Rev.) Philadelphia: W. B. Saunders, 1943. Pp. xii+520.

RAPAPORT, D. Emotions and memory. Baltimore: Williams and Wilkins, 1942. Pp. iii+282.

SHRODES, CAROLINE, VAN GUNDY, J., & HUSBAND, R. W.(Eds.) Psychology through literature. An anthology. New York: Oxford Univ. Press, 1943. Pp. xi+389.

WHITNEY, D. D. Family treasures. Lancaster: Jacques Cattell Press, 1943. Pp. 299.

———, Guide Book 1943–44. (1st Ed.) Washington: National Training School for Boys, 1943. Pp. 58.

NOTES AND NEWS

RODERICK N. MENZIES, on leave of absence from Sarah Lawrence College to the Office of Civilian Defense, died suddenly on Monday, February 22, 1943.

JAMES CARLETON BELL, for the past twenty years professor of education, City College (New York), and founder of *The Journal of Educational Psychology*, retired on January 31.

NORMAN FENTON, professor of education at Stanford University, has been appointed consultant in mental hygiene and delinquency prevention by the California Youth Correction Authority. He was formerly director of the Bureau of Juvenile Research in California.

G. L. FREEMAN of Northwestern University has been granted a leave of absence to take a commission in the United States Naval Reserve.

DOUGLAS PARRY, clinical psychologist, Syracuse University, has been appointed clinical psychologist, Bureau of Child Guidance, Southern Illinois Normal University, Carbondale. In addition to his work with children, Dr. Parry will give a course in guidance procedures, including individual guidance of freshmen and sophomores. REBA HARTLEY, who became a member of the department of physiology and health education last September, was recently appointed to the staff of the Bureau of Child Guidance.

The REVEREND ROBERT E. BRENNAN, O. P., founder of the Thomistic Institute, Providence (R. I.) College, has been appointed visiting professor of psychology, Institute of Psychology, University of Montreal. Father Brennan is giving a course in the history of experimental psychology.

DAEL B. WOLFLE, of the University of Chicago, spoke at a meeting of the Minnesota Chapter of Psi Chi, at the University of Minnesota, Wednesday, January 27, on "Teaching Psychology in Wartime."

RICHARD ARTHUR IBISON has left the Kemper Military School to accept a psychological internship at the Wichita Guidance Center. On the same date, MILDRED THOMAS joined the Center staff as a remedial teacher.

The National Training School for Boys has published an interesting Guide Book for boys entering the institution. DR. STEPHEN HABBE, U. S. Public Health Service Psychologist, is responsible for the preparation of the Guidebook. Psychologists who are interested may secure a copy by

writing to the National Training School, Washington, D. C., on the stationery of their organization.

A new monograph series, entitled *Applied Psychology Monographs*, has been launched by the American Association for Applied Psychology. The publisher is Stanford University Press. Inquiries concerning the new series, and manuscripts for publication should be addressed to the editor, DR. H. S. CONRAD, College Entrance Examination Board, Princeton, N. J.

Books For Prisoners of War. A request has been received from the War Prisoners' Aid of the Young Men's Christian Associations for books in the field of psychology, to be sent to allied prisoners of war who are in Axis hands. In the prisoner of war camps, there are teachers and graduate students in psychology, as well as classes organized by the prisoners themselves; requests for psychology textbooks and treatises are therefore frequent, and their provision is a great lift to morale. War Prisoners' Aid has had success in getting books into the hands of prisoners of war and from time to time prepares shipments to be sent across the seas.

One difficulty in transmitting books is that before shipment the approval of the censor of each book has to be obtained. Hence, a general request for donations to which, without doubt, many psychologists would respond, would be of little value. The procedure now worked out is that of requesting each psychologist to make up a list of the books that he is willing to donate for this worthy purpose and send it to MR. PAUL B. ANDERSON, War Prisoners' Aid, World's Committee of the Y.M.C.A., 347 Madison Avenue, New York City, who will check the lists with the censor and write to the prospective donor informing him which books will be accepted. Books sent to prisoners of war may not have any marks or erasures in them and the date of publication may not be later than September 1, 1939.

It is hoped that many psychologists will respond to this appeal by sending books within their own field. There have been requests from prisoners of war for technical and professional books, as well as for more popular types of books which are made available through other than professional sources.

The RESEARCH COUNCIL ON PROBLEMS OF ALCOHOL has announced an award of $1,000 for "outstanding research on alcoholism during 1943." The work must contribute new knowledge in some branch of medicine, biology or sociology important to the understanding or prevention or treatment of alcoholism. Citizens of the United States, Canada or Latin America are eligible for the award.

The project may have been inaugurated at any time in the past or during the year 1943, providided: (a) that a substantial part of the work be carried on during the year 1943; (b) that it be developed to a point at which significant conclusions are possible before the end of the year; and (c) that a report on the work has not been previously announced and described before a scientific body or previously published. It is desirable, *but not necessary*, that those planning to work for the award send to the

Council before March 1, 1943, a statement of such intention. A report of the work and resulting conclusions must be submitted to the Research Council on Problems of Alcohol on or before February 15, 1944.

The Committee of Award will consist of five members: an officer of the American Association for the Advancement of Science, and four representatives of the Scientific Committee of the Research Council on Problems of Alcohol.

If the committee is not convinced of the outstanding merit of the research done during 1943, as described in reports submitted, it may, at its discretion, postpone the award for another year, or until such time as work of such merit has been performed.